

THIS WEEK

EDITORIALS

WORLD VIEW Obama's Sputnik moment could crash **p.137**



HAGFISH Carcass scavenger absorbs nutrients through its skin **p.138**

FLEXIBLE FRIEND Organic solar cells stretch the future of clean energy **p.139**

Embrace change

US biomedical scientists should support bold plans to transform the process of drug development. Now is not the time for disunity.

At an open meeting at the US National Institutes of Health (NIH) next Monday, agency director Francis Collins will outline his timely vision for a new centre dedicated to translational medicine and therapeutics. Collins sees the new National Center for Advancing Translational Sciences (NCATS) as an incubator for potentially promising nascent therapeutics, to help them through the drug-development pipeline to the point where a risk-averse industry will license them and bring them to the clinic. He wants it to recharge a dwindling cadre of experts in clinical pharmacology. And he sees it as a way of reinventing the woefully inefficient drug-development process itself. But not everyone at Monday's meeting will share Collins's enthusiasm for the project. Establishing the centre will entail breaking up the NIH's National Center for Research Resources (NCRR). More than 1,200 comments, mainly from NCRR constituents fearful for the future of their programmes, have poured into an NIH feedback website.

The process to set up the centre, with an initial budget of more than US\$600 million, has certainly been hasty, and has by its nature alienated a significant element of the NIH-funded community. But Collins is right to seek to accomplish quickly what otherwise threatens to become a drawn-out and even more disruptive period of necessary change. Band-Aids are best torn off quickly.

Dissent has largely focused on how Collins plans to make that change. The largest piece of the proposed new centre, the \$490-million Clinical and Translational Science Awards, will be transferred from the \$1.3-billion NCRR. In what can only be termed an executive decision, taken quickly in December and presented to the NIH community effectively as a fait accompli, Collins wants to dismantle the remaining 60% of the NCRR and move its programmes elsewhere — for example, to the National Institute of General Medical Sciences, the National Institute of Biomedical Imaging and Bioengineering, and to his own office. There, an 'infrastructure entity' will, among other things, administer primate and non-primate disease-model research resources, grants for shared and high-end instrumentation, and training and career development for animal medicine. There have been many calls for a slower, more considered and analytical process to decide the future of the NCRR, including a letter from 16 senators concerned for a \$229-million NCRR programme that supports biomedical infrastructure in mainly rural states that have historically received relatively little NIH money.

It is certainly risky to dismember a \$1.3-billion centre and scatter its programmes without the months of analysis that, for instance, went into the decision to create a new addictions institute at the NIH in 2013. And grant recipients of NCRR-supported programmes have been forced to settle for verbal promises from Collins that the money, staff and commitment will remain unchanged in their new institutional homes. In dire budget times for the agency, such assurances are understandably cold comfort. The new NCATS will also take tremendous institutional investment to make it work, far more than is reducible to dollars and cents. The learning curve for all involved will be steep.

But Collins was hired to lead, and leading he is. What's more, he has reason to hurry. It is entirely possible that he will be out of a job 18 months from now if Republicans capture the White House in next year's presidential election. It's worth noting, too, that the new

Collins is right to seek to accomplish quickly what otherwise threatens to become a drawn-out period of necessary change.

'infrastructure entity' in his office will be overseen by James Anderson, a thoughtful manager with a reputation for being smart, effective and organized — and with a keen eye for checking that programmes are well run. The NIH should certainly make sure in two or three years that, whoever is in charge, Collins's promise to protect NCRR programmes has been kept — provided they are performing.

But the arguments about the NCRR's dissolution must stop here. They risk becoming a distraction and, what's worse, a political liability on Capitol Hill, where lawmakers with their knives out for expendable programmes may find disunity an invitation. The real driver of the angst about the future of NCRR programmes is the precarious position of NIH funding in 2011 and beyond. All those involved should turn their energies to doing everything possible to secure its budget in these extraordinarily difficult times. ■

Notes on a scandal

Events this month have shown that government stances on academic misbehaviour differ wildly.

How an organism is affected by a particular gene mutation, as every geneticist knows, depends on that organism's genetic background. Although an obesity mutation introduced into one strain of mouse might produce a fat animal with diabetes, the same mutation in a mouse strain of slightly different genetic background could create a fat but otherwise healthy animal.

Similarly, the effects of a cry of academic distress seem to depend on a community's societal background. How else to explain the contrasting results of two academic revelations: the plagiarism affair that consumed Germany for two weeks until academic disapproval forced the resignation of the defence minister, Karl-Theodor zu Guttenberg, on 1 March — and an exposé of comparable wrongdoing by the Italian minister of education, Mariastella Gelmini, in 2008, which had zero impact.

The German scandal broke on 16 February, when the daily newspaper *Süddeutsche Zeitung* revealed that the hugely popular Guttenberg had apparently taken a short cut to his doctorate in law by copying other

published works without attribution in his thesis. The report sparked an intense reaction hard to imagine in countries such as the United States and Britain, where the academic achievements (if any) or failures of politicians are not considered serious issues.

German citizens looked to the Internet to discover the extent of Guttenberg's plagiarism, which turned out to be quite shameless. The University of Bayreuth withdrew his PhD and is now investigating whether he had just been careless or had intended to deceive. At first, Guttenberg attempted to underplay the importance of "inadequate footnotes" in a thesis; the issue faded to insignificance, he implied, next to his momentous political mission of reorganizing the German armed forces and controlling their presence in Afghanistan. His popularity among the general public remained undiminished, and Chancellor Angela Merkel, herself a PhD physicist, tried to limit damage to her government by saying that she had "hired a politician, not a scientific assistant". That was a fatal mistake. Within days, tens of thousands of PhD holders had signed a letter deploring her "mockery" of an academic system that represented decency, honour and responsibility — attributes that they insisted should be reflected in a democratic government. Crushed by this attack of righteousness, Guttenberg finally resigned.

Like Guttenberg, Gelmini was a graduate in law. And like him, she felt that her driving ambition justified taking short cuts in academic procedures to get the degree that would help her political career. In 2001 she travelled from her home town of Brescia in the north of Italy to Reggio Calabria, in the far south, to sit her bar exams. At the time, pass rates in the north were below 10%, compared with a rate of suspiciously more than 90% in Reggio Calabria, a city otherwise known for low academic standards. After the press revealed the Reggio Calabria bar exam to be a scam, the Italian academic community called for

Gelmini's resignation — to no avail. The irony of having a minister with responsibility for universities who herself cheerfully admits to having dodged academic rules is not lost on the community.

In Germany, Italy and neighbouring countries in Europe, politicians are frequently drawn from academia. Credentials help political careers,

"It is surprising and gratifying to find that rage against an academic cheat can provoke serious consequences."

and nearly 20% of the German parliament hold PhDs. But then, almost 9% of Italian parliamentarians are university professors, so the differing reactions to calls for resignation prompted by scholastic misdemeanours cannot be down to ignorance about how universities work. Instead, the difference seems to be based on how large a threat each government considers the weapon of moral correctness to be — and how dangerous is the

academic community wielding that weapon.

Should anyone really have expected the government of Silvio Berlusconi to fear such a weapon?

It is more surprising, and gratifying, to find that in Germany, one of the world's richest and most powerful countries, rage against an academic cheat can provoke serious consequences. Not only was Guttenberg popular, but he hadn't previously made any serious political errors that would have seen charges of plagiarism considered the last straw.

Still, there may not be a lesson for many other countries here. Germany is known as the 'country of poets and philosophers' — a rare societal background, and one apparently conducive to propagation of honourable academic values. Like our more fortunate mutant mouse, all there seems plump and healthy, even as it remains unfathomably mysterious to those on the outside. ■

Over the limit

Evidence should be considered when setting policy, but not to the exclusion of other factors.

The evidence is clear. Statistics from the United States, Europe, Asia and Africa all point in the same direction: male drivers are more likely to crash their cars than females. Aggressive behaviour, rule-breaking tendencies and a greater willingness to take risks are all thought to contribute. Taken together, male drivers are a riskier bet, and face higher premiums for car insurance as a result.

Last week, the Court of Justice of the European Communities took a wrecking ball to this seemingly evidence-based policy. From December next year, insurance companies will no longer be able to discriminate on the basis of sex: men should see their premiums fall and women will pay more. The decision was greeted with howls of outrage in some quarters. From newspaper headlines declaring it to be "madness" to radio phone-ins that bemoaned a "lack of common sense", the underlying message from critics was that discrimination against men by insurance companies was fair because it was based on evidence.

The court's decision affects more than car insurance — life assurance premiums paid by men could rise to match those paid by women, where previously they were discounted because men on average die earlier.

Nature is a vocal and staunch supporter of evidence-based policy-making. Yet it is important to distinguish between policies that ignore the evidence and those that consider it but do not give it the deciding vote. Although the European ruling flies in the face of available research on accident and death rates, policy-makers have to consider other relevant factors as well as the scientific data. Europe introduced a laudable law in 2004 to ensure equality between men and women in access to and provision of goods and services. (Last week's ruling was prompted by a challenge under human-rights legislation to a get-out clause in the

2004 law for the insurance industry.) In this case, the drive in Europe to reverse practices that survive as a legacy of centuries of inequality and discrimination against women is right to trump the cold logic of the statistics. Besides, it is far from clear how insurance firms translate the sex differences to quantitative hikes in premiums for young men. The best evidence-based decision-making is also transparent.

Science, and so the data and evidence it gathers, serves society best when it is viewed as part of a wider assessment of risk — not an all-powerful framework within which policy must be placed. A report last week from the UK select committee on science and technology complained that the British government had dropped disruption to air travel by natural disasters from its national risk assessments in 2009, yet saw the eruption of Eyjafjallajökull in Iceland in April 2010 close airports, cancel flights and strand thousands of passengers around the world. The Geological Society in London told the committee that Earth scientists had warned the government of the potential for major disruption for a number of years, but felt that the policy-makers had ignored them. Or perhaps, the policy-makers had heard the warnings, weighed up the costs and benefits, but decided the risk was small enough to do nothing except cross their fingers. It is the job of scientists to raise such concerns, but it is the job of politicians to decide when to ignore them. Although the decisions may be judged with hindsight as being poor policy, they remain evidence-based. Politicians too, however, should not shy away from revealing their reasons.

Overstating the role that scientific evidence should have is most inappropriate when the stakes are highest. For global warming, for instance, the evidence is as clear as the statistical difference between male and female drivers. Yet the policy response must also take into account social, economic and political factors. It is legitimate, if short-sighted, to acknowledge that man-made global warming is real but argue that policies to cut emissions are too expensive to pursue. Dismissing those who take such a position as climate sceptics alongside those who deny the evidence is wrong. When setting policies, there are limits to the role that evidence can have. ■

➔ **NATURE.COM**
To comment online,
click on Editorials at:
go.nature.com/xhunq



Science agencies must bite innovation bullet

Before research can rebuild the US economy it must learn from the prosperous heyday of the military-industrial complex, says Daniel Sarewitz.

In January's State of the Union address, President Barack Obama said that the United States had reached "our generation's Sputnik moment", and to respond to international competition he placed science and innovation at the centre of his policy agenda. Amid extraordinary budget pressures, he has called for a US\$7-billion (11.6%) increase in government spending on research for 2012 to help "rebuild" the economy. More money for science is always welcome, but can it deliver on the president's promises? Not necessarily. The post-Sputnik research enterprise that delivered innovation and prosperity is not the same as the one the President is counting on today.

After the Second World War, the United States was the only major world power with a flourishing scientific and industrial base; the country led the world because it had no competitors. It preserved this advantage during the cold war through the Department of Defense's (DOD's) central role in technology development, and through close and persistent ties between the DOD and private industry. Huge procurement budgets cemented these links, creating early markets for technologies such as computers, jet aeroplanes and satellites. To support this technology base, the DOD invested in emerging fields such as computer science, sub-atomic and solid-state physics, and materials science. Resulting waves of innovation created whole industries that helped to fuel the US economy.

Meanwhile, the main civilian science agencies — the National Institutes of Health (NIH), the National Science Foundation, NASA and the Department of Energy (DOE) — developed roles in training scientists and creating knowledge that accelerated innovation. But such agencies were mere booster rockets for the DOD's main engine of innovation. They lacked, and continue to lack, the attributes that accounted for the military's successes — in particular, its focused mission, enduring ties to the private sector and role as an early customer for advanced technologies.

For decades, the DOD's legacy of innovation and economic growth concealed weaknesses in the civilian agencies, which is why so many people still believe that putting more money into civilian research and development is the panacea for what ails US innovation. Former presidential science adviser John Marburger publicly blew the whistle on this simple-minded notion in 2005, when he noted "how primitive the framework is that we use to evaluate policies and assess strength in science and technology". Partly in response to Marburger's provocation, the National Science Foundation initiated its Science of Science and Innovation Policy programme, with the explicit aim of guiding effective science policy-making by creating a foundation of data, theory, methods and models. This worthy goal carries an uncomfortable implication: that the nation's civilian

research and development enterprise had been built on a foundation of hidden assumptions and unsubstantiated claims.

That foundation is beginning to collapse. In 2006, NASA launched its commercial cargo and crew initiative, which funds the private sector to "develop and demonstrate safe, reliable, and cost-effective space transportation capabilities". The programme thus concedes that NASA cannot mount new missions at affordable costs or within reasonable time-frames. Indeed, in January the agency announced that the \$16 billion and six years allocated by Congress for it to build a new heavy-lift vehicle was insufficient, yet a month earlier, SpaceX Corporation in Hawthorne, California, had completed the first orbit and recovery of a commercial spacecraft, for a total cost of less than \$1 billion. In 2007, Congress, having lost confidence in the DOE, decided that the

country needed a new organization to catalyse innovation in energy technology. The result is the Advanced Research Projects Agency-Energy (ARPA-E), which is designed to "bring a freshness, excitement, and sense of mission to energy research" — virtues apparently absent from the existing department. ARPA-E's focus is on high-risk R&D and collaborations between universities and private firms to move promising technologies into the marketplace (see page 145).

Even the NIH, flagship of the nation's health-research system, is now showing signs of self-doubt. The agency receives half of all US government money spent on civilian research, and owes its powerful political and scientific reputation to the widespread belief that spending billions on cutting-edge basic biomedical science is the best route to better health. Yet public-health

indicators in the United States continue to lag behind those in many other nations. The NIH's leaders hope to turn this around with a new National Center for Advancing Translational Science (see page 135).

These changes signal an uncoordinated but government-wide reaction to an inescapable reality. The civilian research agencies were designed as temples of scientific excellence and technological prowess, but they lack the institutional architecture of the cold-war military-industrial complex, and are ill-structured to create and sustain essential links between knowledge generation, technological innovation and desired social outcomes. It is not a matter of basic versus applied research, but of insular versus integrated approaches. If this is truly our generation's Sputnik moment, it will take more than money. The United States must transform its science enterprise to enhance links between research and its application to national needs. ■

THE DEPARTMENT OF DEFENSE'S LEGACY OF INNOVATION CONCEALED WEAKNESSES IN THE CIVILIAN RESEARCH AGENCIES.

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/gt2xft

Daniel Sarewitz is co-director of the Consortium for Science, Policy and Outcomes at Arizona State University, and is based in Washington DC. e-mail: dsarewitz@gmail.com

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

GENETICS

Prenatal test for Down's syndrome

Fetuses can be screened for Down's syndrome using amniocentesis and other invasive methods, which all carry a risk of inducing miscarriage. A non-invasive test that analyses fetal DNA extracted from the mother's blood might one day replace current methods.

Philippos Patsalis at the Cyprus Institute of Neurology and Genetics in Nicosia and his colleagues used an antibody-based method to examine patterns of fetal DNA methylation on chromosome 21, of which Down's patients have an extra copy.

The researchers chose eight regions in which methylation levels differed most between people with and without Down's. They calculated a score for each case and used this to classify samples as unaffected or as Down's syndrome. The researchers accurately identified all 14 cases of Down's out of a pool of 40 samples.

Nature Med. doi:10.1038/nm.2312 (2011)

ZOOLOGY

Fish that feed through their skin

Hagfish can absorb nutrients through their skin, a trait previously unknown in vertebrates.

The animals often feed by boring into carcasses and

eating their way out of this nutrient-rich environment. Chris Glover at the University of Canterbury in Christchurch, New Zealand, and his colleagues found that samples of skin and gills from Pacific hagfish (*Eptatretus stoutii*; **pictured**) can absorb the amino acids glycine and alanine *in vitro*.

The authors propose that the phylogenetic position of hagfish — which have many primitive traits — suggests that nutrient absorption through the skin may have been widespread in early vertebrates, as it is still among aquatic invertebrates.

Proc. R. Soc. B doi:10.1098/rspb.2010.2784 (2011)

GLACIOLOGY

Glacier grows from below

A large fraction of the thick ice sheet that covers East Antarctica's high interior is formed by liquid water freezing on to the bottom of the ice.

Ice sheets were thought to thicken at the top as accumulating snow is transformed into ice. But a radar survey of the structure of the ice sheet around Antarctica's remote Dome A (**pictured**), conducted between 2007 and 2009, reveals that up to half of the 2,400–3,000-metre-thick ice

package found above valleys in the subglacial Gamburtsev mountain range has been added from below, curving the overlying ice.

The discovery, by Robin Bell at Lamont-Doherty Earth Observatory in Palisades, New York, and her colleagues, could help researchers who hope to find and analyse ancient ice for clues to past climate change.

Science doi:10.1126/science.1200109 (2011)

CHEMISTRY

Where did the xenon go?

Only 10% of all the xenon that was present as a gas in Earth's primordial atmosphere is estimated to exist in that form today, leaving researchers wondering where the rest went. David Brock and Gary Schrobilgen at McMaster University in Hamilton, Canada, provide a possible solution to the mystery.

They report the synthesis of an elusive Xe compound, xenon dioxide (XeO₂). Obtained as a yellow crystal by reacting XeF₄ with water at near-ambient conditions,

XeO₂ shows the spectroscopic features of a square-planar solid. This geometry would let Xe atoms replace some silicon atoms in the lattice of SiO₂, one of the most abundant minerals in the Earth's crust.

J. Am. Chem. Soc. doi:10.1021/ja110618g (2011)

ANIMAL BEHAVIOUR

Sharks swim with direction

Some shark species swim towards quite distant targets, at least some of the time, whereas others seemingly move randomly within a small home range.

By feeding tracking data



R. E. BELL/LAMONT-DOHERTY EARTH OBS.



T. MCHUGH/SPPL

from three shark species into a model of terrestrial animal movement, Yanniss Papastamatiou at the University of Florida in Gainesville and his colleagues found that tiger sharks (*Galeocerdo cuvier*) made 'directed walks' at scales of 6–8 kilometres, and thresher sharks (*Alopias vulpinus*) did so at scales of 400–1,900 metres. Both species have large home ranges.

In contrast, reef sharks (*Carcharhinus melanopterus*) have small home ranges and they showed random, unoriented movement at the scales examined.

The findings could help to predict how sharks will disperse in the face of future challenges such as overfishing and climate change, the authors say.

J. Anim. Ecol. doi:10.1111/j.1365-2656.2011.01815.x (2011)

MATERIALS

Solar cells take a stretch

Electronics made of organic materials such as plastic are appealing because of their flexibility and their ability to be used in biological systems, textiles and moving machine parts. Zhenan Bao and her colleagues at Stanford University in California have made the first intrinsically stretchable organic solar cells.

They created the devices by depositing organic films used in previous organic photovoltaics on a pre-stretched rubber base that they then allowed to compress. This introduced buckling waves in the films that made them stretchable.

The team found that the resulting solar cells remained completely functional when stretched lengthwise by up to 27%. Surprisingly, the cells' performance did not depend on the amount by which they were stretched.

Adv. Mater. doi:10.1002/adma.201004426 (2011)

EVOLUTION

A changing eye on the world

The eyes of most animal groups use one of two types of light-sensing cell — rhabdomeric or ciliary photoreceptors. Scientists had thought that ciliary photoreceptors, which include the rods and cones of the human retina, were unique to the eyes of chordates — animals with a backbone precursor. In non-chordates, this cell type had only ever been found in the brain or other body structures.

Now, Yale Passamanek at the University of Hawaii at Manoa and his colleagues have found evidence of ciliary cells in the eyes of larvae of *Terebratalia transversa*, a member of an ancient group of non-chordates, by showing expression of a key ciliary gene for a light-sensitive protein in the eyes. The discovery could mean that the use of these cells to sense directional light evolved more than once.

EvoDevo doi:10.1186/2041-9139-2-6 (2011)

MICROBIOLOGY

Tagging the TB bacterium

The fight against tuberculosis lacks an important weapon: a sensitive and specific diagnostic probe for the causal bacterium, *Mycobacterium tuberculosis*.

Clifton Barry at the National Institutes of Health in Bethesda, Maryland, Benjamin Davis at the University of Oxford, UK, and their co-workers have got around the problem by exploiting an enzymatic reaction that adds the sugar trehalose to the bacterium's surface lipids.

They synthesized a library of small molecules similar to trehalose and tested their ability to act as substrates in this reaction. One of these molecules, FITC-trehalose, which is fluorescent, was incorporated

COMMUNITY CHOICE

The most viewed papers in science

CANCER

Stressing out cancer cells

HIGHLY READ
on www.cell.com
February 2011

A chemical screen has yielded three compounds that kill cells with abnormal numbers of chromosomes — a common feature of cancer cells.

Angelika Amon and her colleagues at the Massachusetts Institute of Technology in Cambridge looked for compounds that block proliferation of mouse cells with extra chromosomes, but that do not affect normal cells.

They found three such compounds, all of which also triggered programmed cell death in the abnormal cells. Two of the chemicals, called AICAR and 17-AAG, also inhibited the growth of human cancer cells with abnormal chromosomes. AICAR stimulates the protein p53, which regulates cell death, and may also exacerbate the stress that these cells experience as a result of the production of excess protein from their extra chromosomes.

Cell 144, 499–512 (2011)

into the cell envelope and 'lit up' bacteria grown in its presence. FITC-trehalose also labelled *M. tuberculosis* living inside macrophages, a type of immune-system cell it frequently infects. The authors say that this technique could be used to follow how the bacterium infects macrophages.

Nature Chem. Biol. doi:10.1038/nchembio.539 (2011)

ARCHAEOLOGY

Stone Age home by the sea

Some early Americans preferred to stay close to the sea, judging by discoveries on the Channel Islands off the coast of southern California. These ancient hunting grounds have yielded small stone tools as well as abundant remains of fish, seabirds and shellfish.

Jon Erlandson at the University of Oregon in Eugene and his team dated three sites to between about 12,000 and 11,000 years ago. They found bones from geese, cormorants, albatross, fish and marine mammals, along with the shells of mussels, crabs and abalone.

The hunters must have



J. ERLANDSON

reached the islands by boat, but neither their boats nor human burials have been found. However, their projectile points (pictured) look different from the stone tools made by inland peoples from the Clovis culture, and so are unlikely to have derived from that source. The Channel Island sites support the idea that humans settled the Americas while skirting the Pacific coast.

Science 331, 1181–1184 (2011)
For a longer story on this research, see go.nature.com/cczpua

NATURE.COM

For the latest research published by Nature visit:

www.nature.com/latestresearch

SEVEN DAYS

The news in brief

POLICY

Shutdown avoided

With two days to go before funding legislation expired, the US Congress averted a shutdown of government by voting through a continuing resolution on 2 March that funds most science agencies at fiscal year 2010 levels. But battles over cuts are far from over; the new bill expires on 18 March and two weeks may not be enough to bridge the gulf between the Republicans in the House and Democrats in the Senate on a bill to fund the government for the rest of the 2011 fiscal year. See page 144 for more.

China's power plan

China aims to reduce its energy intensity — energy consumption per unit of economic output — by 16–17% by the end of 2015, Premier Wen Jiabao announced last week. To help meet that target, Wen said that the budget plan for the next five years has lowered the economic growth target to around 7% a year. According to China's National Bureau of Statistics, the country has nearly met the 20% target for reducing energy intensity set five years ago. See go.nature.com/g2vbke for more.

Emergency advice

Scientists should have much more input into the 'National Risk Assessment', a classified document that examines all the risks facing the United Kingdom, according to a 2 March report from the House of Commons Science and Technology Committee. It calls for John Beddington, the government's chief scientific adviser, to be moved from the Department for Business, Innovation and Skills to the Cabinet Office, the most central part of the British

government, to facilitate scientific advice in advance of emergencies. See go.nature.com/52hqjb for more.

Spanish protest

More than 2,500 Spanish scientists, including 150 full professors and 4 research-centre directors, have sent a petition to Prime Minister José Luis Rodríguez Zapatero asking for changes to a science and technology law being debated in the country's parliament. The letter, delivered on 22 February, requests among other things that the law include a plan to give researchers the chance to win five-year, tenure-track contracts. The provision was present in previous drafts but then abandoned. See go.nature.com/rurt2w for more.

RESEARCH

EU mega-projects

The European Commission has shortlisted six projects to compete for two €1-billion (US\$1.4-billion) grants under the Future and Emerging Technologies Flagship Initiative to apply information and communication technologies to societal problems. Ideas include cuddly robots to mitigate the loneliness of ageing and a computer simulation of the human brain. Two winners will be selected at the end of the year. See go.nature.com/ukbzd for more.

Glory crash

NASA's Glory mission, intended to study solar radiation and the effects of airborne particles on climate,

work alongside the astronauts on the station. In its 27-year career, *Discovery* made more space flights — 39 — and carried more crew members than any other NASA shuttle. NASA will retire all of its shuttles by the end of the year.



NASA

Shuttle bids farewell to space station

The Space Shuttle *Discovery* departed from the International Space Station for the last time after completing its final mission on 7 March. During the mission, the shuttle delivered Robonaut 2, a humanoid robot designed to

crashed shortly after it launched from Vandenberg Air Force Base in California on 4 March. The protective fairing surrounding the satellite failed to separate from the Taurus XL rocket that carried it, dragging the craft into icy waters near

SOUND BITE

“2011's biggest problem will be food.”

John Beddington, the UK government's chief scientific adviser, at a meeting on agricultural greenhouse-gas emissions.

Antarctica. The failure is the second major loss for NASA's Earth-observation programme in as many years, following the crash of the Orbiting Carbon Observatory — also borne on a Taurus rocket — in 2009. See page 143 for more.

UK health cash

British medical researchers will benefit from a 30% boost to a key translational medicine fund, the country's coalition government announced this week. Starting in April 2012, up to £775 million (US\$1.2 billion) will be awarded over five years via collaborations established by the National Institute for Health Research between universities and the National Health Service for research on new medical treatments. This is a 30% increase on funding over the previous five years of the programme, says the government, which promised last year to increase health-care research spending.

French loan

The French government announced a total of €260 million (US\$360 million) over 10 years for 11 science projects on 4 March. The money is part of a €35-billion stimulus package, revealed by President Nicolas Sarkozy in December 2009, to boost the country's long-term competitiveness.

The 11 schemes include 9 national infrastructure projects for medical and life-sciences research and 2 biotechnology demonstration projects. See go.nature.com/ustnas for more.

PEOPLE



Academy president

Chemist Bai Chunli (pictured) was named the president of the Chinese Academy of Sciences on 28 February, succeeding Lu Yongxiang, who has led the institution for the past 14 years. Bai, the academy's executive vice-president since 2004, is widely credited for the rapid development of nanotechnology in China. He will lead the academy through an ambitious ten-year scheme known as Innovation 2020 (see *Nature* 470, 15; 2011).

Director resigns

Howard Davies stepped down as director of the London School of Economics (LSE) on 4 March, citing the university's connections with Libyan leader Muammar al-Gaddafi.

The LSE has commissioned an external inquiry into its relationship with Libya, including a £1.5-million (US\$2.4-million) donation to the school from the Gaddafi International Charity and Development Foundation in Tripoli. The inquiry will also look into the validity of the PhD thesis written in 2008 by Gaddafi's second son, Saif, while at the LSE.

Fraud conviction

A former biologist at the Massachusetts Institute of Technology (MIT) in Cambridge was convicted of fraud in a US district court in Boston on 3 March. Luk Van Parijs was fired in 2005 for falsifying data. The fraud conviction arises from a 2003 grant application to the National Institutes of Health (NIH), in which Van Parijs falsely claimed that his lab had developed a transgenic mouse. Van Parijs could face jail time, and has agreed to pay US\$61,117 to MIT, which had reimbursed the NIH for money spent from the grant. He will be sentenced on 14 June.

Minister quits

Archaeologist Zahi Hawass, who was appointed Egypt's minister of antiquities in the dying days of the Mubarak regime, said on 3 March that he would not take a position in

COMING UP

9–10 MARCH

The European Union (EU) Competitiveness Council of science and research ministers meets in Brussels, where they will discuss progress towards a common EU patent system.

14–16 MARCH

Second International Conference on the exploration of the Martian moons Phobos and Deimos is held at NASA's Ames Research Center in Moffett Field, California.

go.nature.com/Gufzqa

the new government. Hawass, who also quit as head of the country's Supreme Council of Antiquities, claims that he is leaving because the security services could no longer adequately ensure the protection of archaeological sites and museums under his jurisdiction. Meanwhile, Essam Sharaf, a pro-science reformist, was named the country's new prime minister on 3 March. See go.nature.com/a62e6t for more.

BUSINESS

Biotech buyout

Daiichi Sankyo, a pharmaceutical company based in Tokyo, has agreed to pay up to US\$935 million for Plexxikon, a biotechnology firm in Berkeley, California, it announced on 28 February. Plexxikon has no medicines on the market, but has drawn attention for its lead drug candidate, PLX4032, which has shown promise in patients with advanced melanoma. Daiichi Sankyo will pay \$805 million up front, and as much as \$130 million more if PLX4032 is approved.

► NATURE.COM

For daily news updates see:

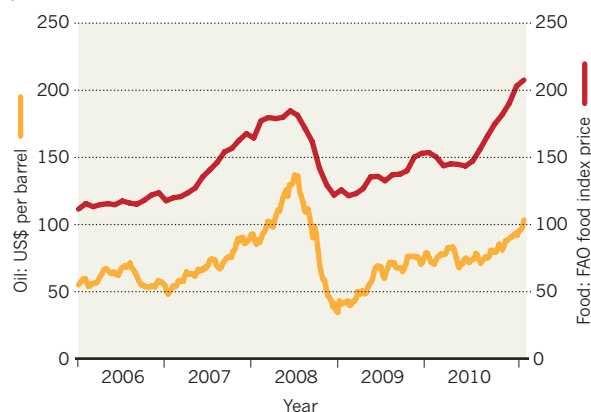
www.nature.com/news

TREND WATCH

Food and oil prices are rising again, after the steep drop that followed the global financial crisis in 2008. The food price index compiled by the Food and Agriculture Organization of the United Nations in Rome has already surpassed its previous record high. The price of crude oil is lagging — it has not yet equalled its 2008 peak of US\$145 a barrel — but it did breach the \$100-a-barrel mark on 2 March. The two are tightly linked; the cost of transportation fuel is one of the main drivers of food prices.

DOUBLE TROUBLE

The prices of food and oil are rising in tandem, heading towards a peak last seen in 2008.



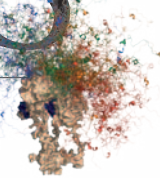
NEWS IN FOCUS

POLICY US budget woes could stifle high-risk energy-research agency **p.145**

SPACE Survey gives a qualified yes to a Mars sample-return mission **p.146**

ENERGY Flawed cable threatens to trip up giant fusion reactor **p.150**

PROTEINS Molecules that function without structure **p.151**



AP PHOTO/A. GALVAN III



Just moments after the launch of the Glory probe, NASA realized that the satellite was doomed.

EARTH SCIENCE

Mourning Glory

NASA satellite crash will hamper solar monitoring and aerosol measurements vital to improving climate models.

BY JEFF TOLLEFSON

The parallels are spooky. The Glory probe was meant to collect crucial climate data as part of NASA's 'A-Train' constellation of Earth-observing satellites, just like the doomed Orbiting Carbon Observatory (OCO) that launched two years ago. Like OCO, Glory took off from Vandenberg Air Force Base in California. And it rode the same model of rocket, the Taurus XL.

Gregory Kopp, a physicist from the University of Colorado in Boulder, who was in charge of a solar-irradiance-monitoring instrument on Glory, couldn't help but feel "apprehensive"

as the probe prepared for countdown on 4 March. Then his worst fears came true.

Minutes after Glory's launch at 2:09 a.m. local time, NASA officials declared the US\$424-million mission a failure. Like OCO, Glory had failed to reach orbit, apparently plunging into the icy southern reaches of the Pacific Ocean. An initial analysis suggests that, just as with the previous mission, the clamshell fairing that protected the satellite during launch had failed to detach, weighing it down.

Glory's failure is another serious blow for NASA's Earth-observation programme. It also raises tough questions for the

company that made the Taurus rocket: Orbital Sciences Corporation, based in Dulles, Virginia. "I understand these things happen with the difficult tasks NASA attempts," Kopp wrote to *Nature* in a remarkably calm e-mail from the Vandenberg base, where he witnessed the failed launch. "I'm of course surprised that after two years of scrutiny, Orbital evidently does not understand what is causing their Taurus launch vehicles to fail — they've now lost three of the last four."

Scientists are also grieving for data that might have been. The loss of Glory's solar monitor increases the risk of a lapse in the continuous 32-year record of the Sun's total energy output, data that are fundamental to climate-change models. It also complicates efforts to compare solar-irradiance data from different instruments.

Kopp had calibrated Glory's Total Irradiance Monitor (TIM) using a ground-based facility at the University of Colorado. The same system had been used to test the Precision Monitoring Sensor (PREMOS) — a Swiss-built instrument launched last June aboard the French satellite PICARD — and to validate the measurements taken by an older version of TIM now flying on NASA's Solar Radiation and Climate Experiment (SORCE) satellite. The system is designed to avoid mismatches of data sets in the solar record, an issue that has been addressed by using different instruments to take overlapping measurements. Glory represented the biggest test yet of this calibration system (see *Nature* **469**, 457–458; 2011).

"It's really sad," says Werner Schmutz, director of the World Radiation Center in Davos, Switzerland, and principal investigator for PREMOS. But Schmutz is fairly confident that one of the existing solar-monitoring satellites will survive until the next sensor goes up; NASA is already planning to launch another instrument designed by Kopp's team in 2015. Even if there is a monitoring hiatus, adds Schmutz, the calibration system should still be able to link future satellite data with previous records.

Glory also carried the Aerosol Polarimetry Sensor, which would have provided climate modellers with important data about dust, sulphate haze, black carbon and other ►

► **NATURE.COM**
For the latest on
climate research,
visit:
nature.com/climate

► particulates in the atmosphere. By analysing the polarization of light bouncing off the particles, scientists would have been able to discern their size, shape and chemical make-up — key information for understanding their effect on clouds and global climate. These effects collectively represent the greatest area of uncertainty in current climate models, so scientists have been talking about putting a sensor like this in space for more than a decade.

“The magnitude of the aerosol forcing is a big determinant of where we are going to go in the future,” says Gavin Schmidt, a climate modeller at NASA’s Goddard Institute for Space Studies in New York. “We have the technology to pin this down more than it is being pinned down, and yet that technology is sitting at the bottom of the ocean.”

Glory’s failure puts NASA in a quandary. OCO, which would have mapped atmospheric carbon dioxide levels with unprecedented accuracy, enabling scientists to improve definitions of carbon sinks and sources, was deemed so important that the agency planned a rapid-replacement mission. OCO-2 is now scheduled to launch in February 2013 — once again aboard a Taurus XL rocket built by Orbital Sciences. Michael Luther, deputy associate administrator for science programmes at NASA, says the agency will complete the investigation into Glory’s failure and then “adjust our plans [for OCO-2] appropriately”.

Engineers had determined that the most likely culprit for OCO’s failure to reach orbit in 2009 was the system for separating the fairing from the rocket body, which relied on explosives to push the fairing’s halves away. After OCO failed, the system was redesigned to use pressurized cold nitrogen gas instead. “We really felt like we had the problem nailed,” said Richard Straka, deputy manager of Orbital Science’s launch-system group, at a NASA press conference shortly after Glory’s launch.

Jonathan McDowell, an astrophysicist at the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, who follows rocket launches closely, suggests that Glory’s failure might come down to a wiring error. He points out that Orbital Science’s Pegasus and Minotaur rockets, which between them have made almost 20 successful launches, use fairings and fairing-release systems similar to those on its Taurus rockets. And because the Taurus had five successful launches between 1994 and 2000 before the string of failures began, McDowell sees no evidence of a fundamental design problem. “So it’s either sloppiness ... or really bad luck.” ■

Additional reporting by Geoff Brumfiel.

POLICY

US scientists in budget limbo

Researchers face anxious wait as negotiations continue in Congress over 2011 budget.

BY EUGENIE SAMUEL REICH

The US Congress narrowly avoided a government shutdown last week by passing a continuing resolution to fund federal activities until 18 March. But the two-week reprieve is prolonging scientists’ anxiety over the final 2011 budget that may emerge from negotiations between the Republican-majority House and the Democrat-majority Senate. The delay is also raising fears about how drastic the cuts to science might be.

“It’s a time of great uncertainty and the [scientific] community is very concerned,” says John Marburger, vice-president for research at the State University of New York at Stony Brook. At his university, researchers continue to put in proposals for federal funding, but they worry that a tighter budget will reduce success rates. Marburger, who was science adviser to former president George W. Bush, has spoken out against the cuts proposed by House Republicans. Post-docs and graduate students, who, he estimates, make up 80% of researchers supported by federal grants, will be hit hardest. “It will put people out on the streets.”

The latest continuing resolution cuts US\$4 billion from federal budgets, although it largely spares science. And if a 2011 budget is eventually agreed, it is unlikely to contain the full \$61 billion in cuts passed by the House on

19 February, which are not expected to pass the Senate (see ‘Budget ping pong’). But House Republicans have said that these proposals will form the basis for future negotiations. Judging by an alternative bill put forward by Senate Democrats on 4 March, the final budget will undoubtedly be lower than 2010 levels.

The uncertainty can be paralyzing, says Jennifer Zeitzer, director of legislative affairs for the Federation of American Societies for Experimental Biology (FASEB) in Bethesda, Maryland. “The longer the National Institutes of Health (NIH) is under a continuing resolution, the more people may have questions about their grants,” she says. The latest continuing resolution “just kicks the can down the road”. Neurobiologist Darcy Kelley of Columbia University in New York agrees that the mere threat of cuts may be enough to hamper scientists’ careers. “With funding levels certain to drop quite drastically, I think people would be even more cautious about hiring somebody new in the lab,” she says.

If Congress can’t reach agreement in the next two weeks, it will again risk a government shutdown. William Talman, president of FASEB and a neuroscientist at both the University of Iowa and the Veterans Affairs Medical Center in Iowa City, vividly remembers the last government shutdown that began in December 1995. As a physician, Talman was considered to be ‘essential personnel’ and was funded to care for patients. But his research — a study of neural regulation of blood pressure in rats,

“It will put people out on the streets.”

BUDGET PING PONG

House Republicans and Senate Democrats are battling over federal agencies’ 2011 budgets.

Agency	House proposed cut (US\$ million)	% change from 2010	Senate proposed cut (US\$ million)	% change from 2010
Environmental Protection Agency	2,859	–28%	398	–4%
Centers for Disease Control and Prevention	1,397	–22%	173 increase	+3%
Department of Energy’s Office of Science	893.2	–18%	162	–3%
National Institute of Standards and Technology	160	–19%	108	–13%
NASA	303	–2%	224	–1%
National Institutes of Health	1,629	–5%	86	–0.3%
National Science Foundation	360	–5%	73	–1%

SOURCE: US HOUSE AND SENATE APPROPRIATIONS COMMITTEES

► particulates in the atmosphere. By analysing the polarization of light bouncing off the particles, scientists would have been able to discern their size, shape and chemical make-up — key information for understanding their effect on clouds and global climate. These effects collectively represent the greatest area of uncertainty in current climate models, so scientists have been talking about putting a sensor like this in space for more than a decade.

“The magnitude of the aerosol forcing is a big determinant of where we are going to go in the future,” says Gavin Schmidt, a climate modeller at NASA’s Goddard Institute for Space Studies in New York. “We have the technology to pin this down more than it is being pinned down, and yet that technology is sitting at the bottom of the ocean.”

Glory’s failure puts NASA in a quandary. OCO, which would have mapped atmospheric carbon dioxide levels with unprecedented accuracy, enabling scientists to improve definitions of carbon sinks and sources, was deemed so important that the agency planned a rapid-replacement mission. OCO-2 is now scheduled to launch in February 2013 — once again aboard a Taurus XL rocket built by Orbital Sciences. Michael Luther, deputy associate administrator for science programmes at NASA, says the agency will complete the investigation into Glory’s failure and then “adjust our plans [for OCO-2] appropriately”.

Engineers had determined that the most likely culprit for OCO’s failure to reach orbit in 2009 was the system for separating the fairing from the rocket body, which relied on explosives to push the fairing’s halves away. After OCO failed, the system was redesigned to use pressurized cold nitrogen gas instead. “We really felt like we had the problem nailed,” said Richard Straka, deputy manager of Orbital Science’s launch-system group, at a NASA press conference shortly after Glory’s launch.

Jonathan McDowell, an astrophysicist at the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, who follows rocket launches closely, suggests that Glory’s failure might come down to a wiring error. He points out that Orbital Science’s Pegasus and Minotaur rockets, which between them have made almost 20 successful launches, use fairings and fairing-release systems similar to those on its Taurus rockets. And because the Taurus had five successful launches between 1994 and 2000 before the string of failures began, McDowell sees no evidence of a fundamental design problem. “So it’s either sloppiness ... or really bad luck.” ■

Additional reporting by Geoff Brumfiel.

POLICY

US scientists in budget limbo

Researchers face anxious wait as negotiations continue in Congress over 2011 budget.

BY EUGENIE SAMUEL REICH

The US Congress narrowly avoided a government shutdown last week by passing a continuing resolution to fund federal activities until 18 March. But the two-week reprieve is prolonging scientists’ anxiety over the final 2011 budget that may emerge from negotiations between the Republican-majority House and the Democrat-majority Senate. The delay is also raising fears about how drastic the cuts to science might be.

“It’s a time of great uncertainty and the [scientific] community is very concerned,” says John Marburger, vice-president for research at the State University of New York at Stony Brook. At his university, researchers continue to put in proposals for federal funding, but they worry that a tighter budget will reduce success rates. Marburger, who was science adviser to former president George W. Bush, has spoken out against the cuts proposed by House Republicans. Post-docs and graduate students, who, he estimates, make up 80% of researchers supported by federal grants, will be hit hardest. “It will put people out on the streets.”

The latest continuing resolution cuts US\$4 billion from federal budgets, although it largely spares science. And if a 2011 budget is eventually agreed, it is unlikely to contain the full \$61 billion in cuts passed by the House on

19 February, which are not expected to pass the Senate (see ‘Budget ping pong’). But House Republicans have said that these proposals will form the basis for future negotiations. Judging by an alternative bill put forward by Senate Democrats on 4 March, the final budget will undoubtedly be lower than 2010 levels.

The uncertainty can be paralyzing, says Jennifer Zeitzer, director of legislative affairs for the Federation of American Societies for Experimental Biology (FASEB) in Bethesda, Maryland. “The longer the National Institutes of Health (NIH) is under a continuing resolution, the more people may have questions about their grants,” she says. The latest continuing resolution “just kicks the can down the road”. Neurobiologist Darcy Kelley of Columbia University in New York agrees that the mere threat of cuts may be enough to hamper scientists’ careers. “With funding levels certain to drop quite drastically, I think people would be even more cautious about hiring somebody new in the lab,” she says.

If Congress can’t reach agreement in the next two weeks, it will again risk a government shutdown. William Talman, president of FASEB and a neuroscientist at both the University of Iowa and the Veterans Affairs Medical Center in Iowa City, vividly remembers the last government shutdown that began in December 1995. As a physician, Talman was considered to be ‘essential personnel’ and was funded to care for patients. But his research — a study of neural regulation of blood pressure in rats,

“It will put people out on the streets.”

BUDGET PING PONG

House Republicans and Senate Democrats are battling over federal agencies’ 2011 budgets.

Agency	House proposed cut (US\$ million)	% change from 2010	Senate proposed cut (US\$ million)	% change from 2010
Environmental Protection Agency	2,859	–28%	398	–4%
Centers for Disease Control and Prevention	1,397	–22%	173 increase	+3%
Department of Energy’s Office of Science	893.2	–18%	162	–3%
National Institute of Standards and Technology	160	–19%	108	–13%
NASA	303	–2%	224	–1%
National Institutes of Health	1,629	–5%	86	–0.3%
National Science Foundation	360	–5%	73	–1%

SOURCE: US HOUSE AND SENATE APPROPRIATIONS COMMITTEES

funded by the NIH — came to a grinding halt, because his lab was then located in the government-funded Veterans Affairs Medical Center. “In the labs there was no engineering support, no mechanical support, nothing considered non-essential,” he says. “Essentially the labs were in lockdown.”

Another temporary funding measure could avert that outcome, but would simply prolong the uncertainty. At the Thomas Jefferson National Accelerator Facility in Newport News, Virginia, a temporary shut-down to upgrade the particle accelerator is scheduled for this year. Robert McKeown, the facility's deputy director for science, had been planning to hire extra people to work on the upgrade, to keep the project on track. But

“if the funding situation continues two weeks at a time, I’m afraid we won’t be able to make decisions to hire people”.

The delay to a finalized 2011 budget also means that any cuts will feel more dramatic when they come. A proposed 18% cut to the Department of Energy’s Office of Science 2011 budget would require a cut of 30% over the remaining seven months of the fiscal year. At the Thomas Jefferson facility, that would mean the suspension of the accelerator upgrade, McKeown says, and lay-offs for 300 of the roughly 800 staff members working on the site.

Astronomer Scott Tremaine of the Institute for Advanced Study in Princeton, New Jersey — who advised on the National Academy of

Sciences’ decadal survey of astronomy and astrophysics priorities in 2010 — says that the uncertainty highlights a major problem of the US system. Budgets are set year-to-year — or, at the moment, fortnight-to-fortnight — through congressional negotiations, so long-term, international collaborations are difficult to plan, and assumptions about future budgets are liable to be wrong. For example, the decadal survey worked with figures from NASA projecting a flat budget over the coming decade, and with figures from the National Science Foundation that assumed a doubling of funds relative to a 2007 baseline. Now, says Tremaine, “that is looking unrealistic”. ■

Additional reporting by Meredith Wadman.

POLICY

Risky energy research faces uncertain future

ARPA-E’s aggressive approach to managing research wins support — but perhaps not federal dollars.

BY JEFF TOLLEFSON

When the US energy department’s new agency for high-risk, high-impact energy research announced a competition for carbon-capture technologies in December 2009, David Sholl decided to take the plunge. Sholl and his colleagues at the Georgia Institute of Technology in Atlanta had an idea for membranes that make use of a new class of porous crystalline compound to filter carbon dioxide from power-plant exhaust. A month later they had submitted a proposal to the agency, called ARPA-E (Advanced Research

Projects Agency-Energy), and in April 2010 they won a US\$1-million, two-year award.

The first money arrived in July, and two weeks later, Sholl says, his team got a lesson in ARPA-E’s hands-on style. “They started calling and saying, ‘What have you got done?’” Sholl says. “I’m used to the money-over-the-fence approach with science funding, but this is a much more collaborative sort of relationship.” A year into the first round of projects funded by ARPA-E, scientists, businesses and venture capitalists say that the model is already creating a powerful ecosystem that cultivates entrepreneurial science. At least six ARPA-E projects,

targeting technologies from solar power to wind energy and batteries, have already gone on to earn additional backing from venture capitalists, and President Barack Obama’s administration is eager to scale up the programme. But budget cutting in Congress could stifle the agency before it has a chance to prove itself.

ARPA-E, which received its first funding in 2009, was modelled on the Defense Advanced Research Projects Agency (DARPA), famous for its development of the Internet. Like DARPA, ARPA-E seeks to identify game-changing ideas that are too radical for agencies such as the National Science Foundation. Its research managers then actively cultivate each project, unlike the hands-off approach taken by basic-research agencies.

The approach leaves little room for serendipity, but Sholl says that it is an effective way to keep research moving and focused on a specific, commercially relevant goal. “One of the huge advantages of the ARPA-E programme is that there is a well-defined industrial target,” he says. “ARPA-E is kind of god’s gift to venture capital,” agrees Matthew Nordan, vice-president of the venture-capital firm Venrock in Cambridge, Massachusetts. “It ▶

ENERGY INVESTMENT

ARPA-E has rolled out grants for more than 100 projects in the past year and a half, ranging from electrofuels (transportation fuels produced from carbon dioxide) to energy-storage technologies for the grid.

\$365.5 million

Total ARPA-E funding:
121 projects

\$158.4 million

General*:
41 projects

*Including: energy storage, \$33.1 million; renewable power, \$28.5 million; biomass energy, \$27.6 million; vehicle technologies, \$19.5 million; direct solar fuels, \$15.9 million; building efficiency, \$15.1 million; carbon capture, \$11.2 million; waste-heat capture, \$4.7 million; water, \$1.8 million; conventional energy, \$1 million

\$44.5 m

Electrofuels:
13 projects

\$36.3 m

Batteries for Electrical
Energy Storage in
Transportation (BEEST):
10 projects

\$32 m

Innovative Materials
and Processes for
Advanced Carbon Capture
Technologies (IMPACCT):
15 projects

\$36.3 m

Agile Delivery of Electrical
Power Technology (ADEPT):
14 projects

\$30.3 m

Building Energy Efficiency
Through Innovative
Thermoelectrics (BEETIT):
16 projects

\$27.7 m

Grid-Scale Rampable
Intermittent Dispatchable
Storage (GRIDS):
12 projects

funded by the NIH — came to a grinding halt, because his lab was then located in the government-funded Veterans Affairs Medical Center. “In the labs there was no engineering support, no mechanical support, nothing considered non-essential,” he says. “Essentially the labs were in lockdown.”

Another temporary funding measure could avert that outcome, but would simply prolong the uncertainty. At the Thomas Jefferson National Accelerator Facility in Newport News, Virginia, a temporary shut-down to upgrade the particle accelerator is scheduled for this year. Robert McKeown, the facility's deputy director for science, had been planning to hire extra people to work on the upgrade, to keep the project on track. But

“if the funding situation continues two weeks at a time, I’m afraid we won’t be able to make decisions to hire people”.

The delay to a finalized 2011 budget also means that any cuts will feel more dramatic when they come. A proposed 18% cut to the Department of Energy’s Office of Science 2011 budget would require a cut of 30% over the remaining seven months of the fiscal year. At the Thomas Jefferson facility, that would mean the suspension of the accelerator upgrade, McKeown says, and lay-offs for 300 of the roughly 800 staff members working on the site.

Astronomer Scott Tremaine of the Institute for Advanced Study in Princeton, New Jersey — who advised on the National Academy of

Sciences’ decadal survey of astronomy and astrophysics priorities in 2010 — says that the uncertainty highlights a major problem of the US system. Budgets are set year-to-year — or, at the moment, fortnight-to-fortnight — through congressional negotiations, so long-term, international collaborations are difficult to plan, and assumptions about future budgets are liable to be wrong. For example, the decadal survey worked with figures from NASA projecting a flat budget over the coming decade, and with figures from the National Science Foundation that assumed a doubling of funds relative to a 2007 baseline. Now, says Tremaine, “that is looking unrealistic”. ■

Additional reporting by Meredith Wadman.

POLICY

Risky energy research faces uncertain future

ARPA-E’s aggressive approach to managing research wins support — but perhaps not federal dollars.

BY JEFF TOLLEFSON

When the US energy department’s new agency for high-risk, high-impact energy research announced a competition for carbon-capture technologies in December 2009, David Sholl decided to take the plunge. Sholl and his colleagues at the Georgia Institute of Technology in Atlanta had an idea for membranes that make use of a new class of porous crystalline compound to filter carbon dioxide from power-plant exhaust. A month later they had submitted a proposal to the agency, called ARPA-E (Advanced Research

Projects Agency-Energy), and in April 2010 they won a US\$1-million, two-year award.

The first money arrived in July, and two weeks later, Sholl says, his team got a lesson in ARPA-E’s hands-on style. “They started calling and saying, ‘What have you got done?’” Sholl says. “I’m used to the money-over-the-fence approach with science funding, but this is a much more collaborative sort of relationship.” A year into the first round of projects funded by ARPA-E, scientists, businesses and venture capitalists say that the model is already creating a powerful ecosystem that cultivates entrepreneurial science. At least six ARPA-E projects,

targeting technologies from solar power to wind energy and batteries, have already gone on to earn additional backing from venture capitalists, and President Barack Obama’s administration is eager to scale up the programme. But budget cutting in Congress could stifle the agency before it has a chance to prove itself.

ARPA-E, which received its first funding in 2009, was modelled on the Defense Advanced Research Projects Agency (DARPA), famous for its development of the Internet. Like DARPA, ARPA-E seeks to identify game-changing ideas that are too radical for agencies such as the National Science Foundation. Its research managers then actively cultivate each project, unlike the hands-off approach taken by basic-research agencies.

The approach leaves little room for serendipity, but Sholl says that it is an effective way to keep research moving and focused on a specific, commercially relevant goal. “One of the huge advantages of the ARPA-E programme is that there is a well-defined industrial target,” he says. “ARPA-E is kind of god’s gift to venture capital,” agrees Matthew Nordan, vice-president of the venture-capital firm Venrock in Cambridge, Massachusetts. “It ▶

ENERGY INVESTMENT

ARPA-E has rolled out grants for more than 100 projects in the past year and a half, ranging from electrofuels (transportation fuels produced from carbon dioxide) to energy-storage technologies for the grid.

\$365.5 million

Total ARPA-E funding:
121 projects

\$158.4 million

General*:
41 projects

*Including: energy storage, \$33.1 million; renewable power, \$28.5 million; biomass energy, \$27.6 million; vehicle technologies, \$19.5 million; direct solar fuels, \$15.9 million; building efficiency, \$15.1 million; carbon capture, \$11.2 million; waste-heat capture, \$4.7 million; water, \$1.8 million; conventional energy, \$1 million

\$44.5 m

Electrofuels:
13 projects

\$36.3 m

Batteries for Electrical
Energy Storage in
Transportation (BEEST):
10 projects

\$32 m

Innovative Materials
and Processes for
Advanced Carbon Capture
Technologies (IMPACCT):
15 projects

\$36.3 m

Agile Delivery of Electrical
Power Technology (ADEPT):
14 projects

\$30.3 m

Building Energy Efficiency
Through Innovative
Thermoelectrics (BEETIT):
16 projects

\$27.7 m

Grid-Scale Rampable
Intermittent Dispatchable
Storage (GRIDS):
12 projects

► provides an independent yardstick as to whether something is interesting enough and whether it meets its yard posts.”

At ARPA-E's second annual Energy Innovation Summit in Washington DC last week, the mood was upbeat and the speeches were inspirational, underscoring science's role in keeping the United States competitive in the global race for clean energy. But the fiscal backdrop is ominous. In Congress, Democrats and Republicans remain at loggerheads over federal spending, having temporarily averted a government shutdown last week when the House of Representatives passed a resolution that funds the government until 18 March (see page 144). Although the White House proposed a budget of \$300 million for ARPA-E in the current fiscal year, the continuing resolution contains just \$50 million for the agency, enough to launch at most one new research programme.

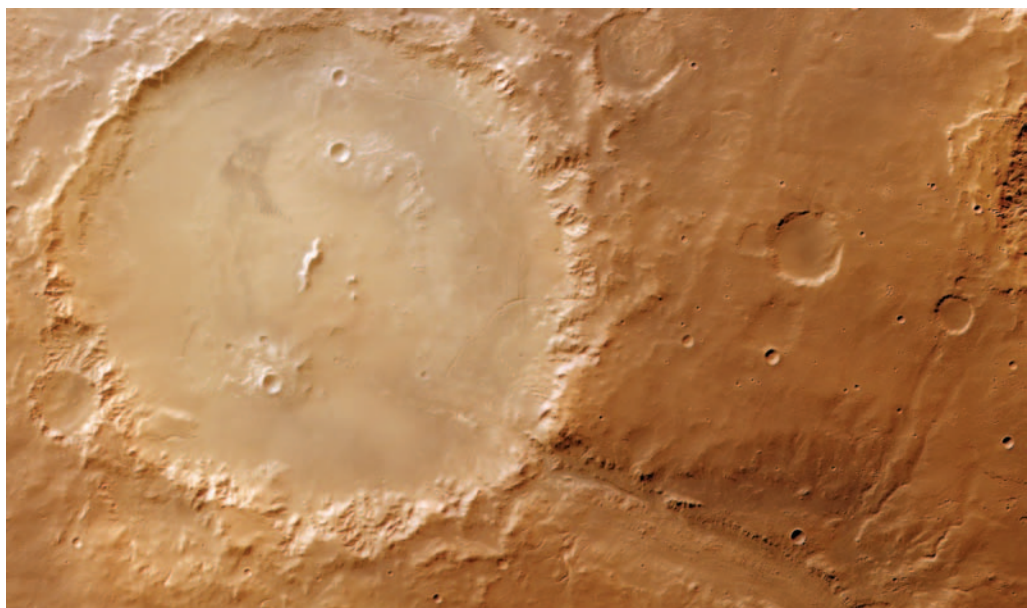
To date, the agency has committed \$365.5 million to 121 projects in various fields (see 'Energy investment'). But all of that money came from the original \$400-million appropriation in the economic stimulus package enacted in April 2009. Depending on how much more money comes through this year, the agency is already considering several new programmes in areas such as solar photovoltaics, the conversion of natural gas into liquid fuels, and energy technologies that reduce the consumption of water, rare-earth metals and other crucial materials. For fiscal year 2012, the administration has requested \$550 million, which should allow further expansion.

ARPA-E director Arun Majumdar says that the agency is ready to move ahead once Congress resolves its funding for this year and next. But he received mixed messages from two Republican senators who attended the conference, Lisa Murkowski of Alaska and Lamar Alexander of Tennessee. Both lawmakers, while offering their support for ARPA-E, warned that Congress is in full budget-cutting mode.

The agency will soon begin making difficult choices about its first round of projects. Each project has measurable milestones, and Majumdar says that the agency will put each one on green, yellow and red alerts depending on how much progress it has made. Project managers will continue to work with the scientists to help them meet their targets, but yellow and red signal trouble if researchers don't start to make progress.

“We will help you as much as we can, but if it doesn't work, it doesn't work,” Majumdar says. “Taxpayer money should not be given to things we know are not going to work.” ■

➔ **NATURE.COM**
For more on clean
energy see:
go.nature.com/e6umpw



Mars's 150-kilometre Holden Crater, which shows signs of a watery past, could be perfect for rock hunting.

PLANETARY SCIENCE

US Mars mission takes pole position

Sample-return trip to go ahead, but only if costs can be cut.

BY ADAM MANN

A showdown over the course of Solar System exploration has ended with a qualified victory for Mars. NASA's planetary-science decadal survey, which sets mission priorities for 2013–22, firmly favours a mission to Mars over a rival one to Jupiter's icy moon Europa (see *Nature* **466**, 168–169; 2010). But the decision marks the beginning of a much bigger battle: to secure the budget to lift the multibillion-dollar project off the survey's pages and into the heavens.

The decadal-survey committee's recommendations, released on 7 March at the Lunar and Planetary Science Conference in Houston, Texas, relied partly on President Barack Obama's 2011 budget request, which projected that NASA's annual planetary-science funding would grow from its current allocation of \$1.36 billion to more than \$1.6 billion by 2015. But Obama's 2012 budget foresees that funding dropping to \$1.2 billion in 2016. On 3 March, planetary-sciences division director James Green told the NASA Advisory Council's science committee that this would create indefinite delays for both the Mars and Europa missions.

“This creates a big gap between what the decadal survey is planning on and what is

available,” agrees Fran Bagenal, a planetary scientist at the University of Colorado, Boulder, and former chairwoman of an external NASA planetary-science advisory committee, who was not involved in creating the report.

The details of the recommendation reflect the committee's attempts to navigate different budget scenarios and maintain a robust research agenda under cash-strapped conditions. “In prioritizing missions, the most important criterion was maximizing science bang per buck,” says Steve Squyres, an astronomer at Cornell University in Ithaca, New York, and chairman of the decadal survey.

The top-ranked flagship mission, the Mars Astrobiology Explorer-Cacher (MAX-C), would use a rover to conduct *in situ* astrobiological experiments, and to collect and store samples for return to Earth. This mission would also deliver the ExoMars rover for the European Space Agency (ESA). “We are at the point in Solar System exploration where what we want to do is beyond the budget of a single nation,” says Wendy Calvin, a geologist at the University of Nevada, Reno, and vice-chair of the decadal survey's Mars panel.

To allow room in the budget for other priorities, the report recommends that the mission should not fly in the next decade

G. NEUKUM, FU-BERLIN/ESA/DLR

► provides an independent yardstick as to whether something is interesting enough and whether it meets its yard posts.”

At ARPA-E's second annual Energy Innovation Summit in Washington DC last week, the mood was upbeat and the speeches were inspirational, underscoring science's role in keeping the United States competitive in the global race for clean energy. But the fiscal backdrop is ominous. In Congress, Democrats and Republicans remain at loggerheads over federal spending, having temporarily averted a government shutdown last week when the House of Representatives passed a resolution that funds the government until 18 March (see page 144). Although the White House proposed a budget of \$300 million for ARPA-E in the current fiscal year, the continuing resolution contains just \$50 million for the agency, enough to launch at most one new research programme.

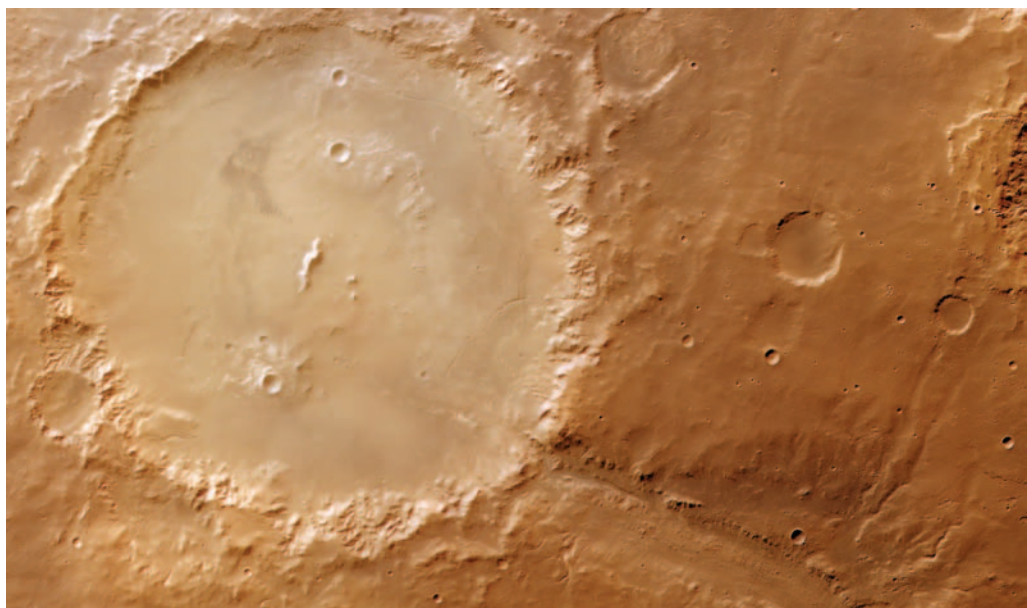
To date, the agency has committed \$365.5 million to 121 projects in various fields (see 'Energy investment'). But all of that money came from the original \$400-million appropriation in the economic stimulus package enacted in April 2009. Depending on how much more money comes through this year, the agency is already considering several new programmes in areas such as solar photovoltaics, the conversion of natural gas into liquid fuels, and energy technologies that reduce the consumption of water, rare-earth metals and other crucial materials. For fiscal year 2012, the administration has requested \$550 million, which should allow further expansion.

ARPA-E director Arun Majumdar says that the agency is ready to move ahead once Congress resolves its funding for this year and next. But he received mixed messages from two Republican senators who attended the conference, Lisa Murkowski of Alaska and Lamar Alexander of Tennessee. Both lawmakers, while offering their support for ARPA-E, warned that Congress is in full budget-cutting mode.

The agency will soon begin making difficult choices about its first round of projects. Each project has measurable milestones, and Majumdar says that the agency will put each one on green, yellow and red alerts depending on how much progress it has made. Project managers will continue to work with the scientists to help them meet their targets, but yellow and red signal trouble if researchers don't start to make progress.

“We will help you as much as we can, but if it doesn't work, it doesn't work,” Majumdar says. “Taxpayer money should not be given to things we know are not going to work.” ■

➔ **NATURE.COM**
For more on clean
energy see:
go.nature.com/e6umpw



Mars's 150-kilometre Holden Crater, which shows signs of a watery past, could be perfect for rock hunting.

PLANETARY SCIENCE

US Mars mission takes pole position

Sample-return trip to go ahead, but only if costs can be cut.

BY ADAM MANN

A showdown over the course of Solar System exploration has ended with a qualified victory for Mars. NASA's planetary-science decadal survey, which sets mission priorities for 2013–22, firmly favours a mission to Mars over a rival one to Jupiter's icy moon Europa (see *Nature* **466**, 168–169; 2010). But the decision marks the beginning of a much bigger battle: to secure the budget to lift the multibillion-dollar project off the survey's pages and into the heavens.

The decadal-survey committee's recommendations, released on 7 March at the Lunar and Planetary Science Conference in Houston, Texas, relied partly on President Barack Obama's 2011 budget request, which projected that NASA's annual planetary-science funding would grow from its current allocation of \$1.36 billion to more than \$1.6 billion by 2015. But Obama's 2012 budget foresees that funding dropping to \$1.2 billion in 2016. On 3 March, planetary-sciences division director James Green told the NASA Advisory Council's science committee that this would create indefinite delays for both the Mars and Europa missions.

“This creates a big gap between what the decadal survey is planning on and what is

available,” agrees Fran Bagenal, a planetary scientist at the University of Colorado, Boulder, and former chairwoman of an external NASA planetary-science advisory committee, who was not involved in creating the report.

The details of the recommendation reflect the committee's attempts to navigate different budget scenarios and maintain a robust research agenda under cash-strapped conditions. “In prioritizing missions, the most important criterion was maximizing science bang per buck,” says Steve Squyres, an astronomer at Cornell University in Ithaca, New York, and chairman of the decadal survey.

The top-ranked flagship mission, the Mars Astrobiology Explorer-Cacher (MAX-C), would use a rover to conduct *in situ* astrobiological experiments, and to collect and store samples for return to Earth. This mission would also deliver the ExoMars rover for the European Space Agency (ESA). “We are at the point in Solar System exploration where what we want to do is beyond the budget of a single nation,” says Wendy Calvin, a geologist at the University of Nevada, Reno, and vice-chair of the decadal survey's Mars panel.

To allow room in the budget for other priorities, the report recommends that the mission should not fly in the next decade

G. NEUKUM, FU BERLIN/ESA/DLR

unless a billion dollars can be shaved off the estimated \$3.5-billion cost to NASA.

The second-choice flagship mission will fly only if its costs can be cut, and if NASA gets a significant budget increase for planetary exploration. Indeed, it is excluded from the panel's 'cost-constrained' mission wish-list (see 'To boldly go ...'). The Jupiter Europa Orbiter (JEO) would map the Jovian moon to assess the extent of the ocean thought to lie beneath its icy surface — a possible habitat for life. But the mission's estimated price tag — \$4.7 billion, adjusted to 2015 dollar values — shocked panel members. "There were a lot of gasps when we saw the bottom line," says Stephen Mackwell, director of the Lunar and Planetary Institute in Houston, Texas, and vice-chair of the inner-planet panel. The high price was a key factor in tipping the decision towards a Mars mission and significantly lengthens the odds against the Europa mission getting off the ground this decade.

The decision to rank the JEO after MAX-C has not met with universal enthusiasm. "We've spent a little too much time on Mars," says Bagenal, adding that the red planet will be visited by the much-delayed and over-budget Mars Science Laboratory rover next year. She also finds it "hard to imagine" that MAX-C can stick to a \$2.5-billion price tag, not least because it would rely on an as-yet-untested landing system.

The third flagship candidate is the Uranus Orbiter and Probe, expected to cost \$2.7 billion. Like the other top recommendations, this project should be delayed or cancelled if costs increase significantly, the committee says. Indeed, these flagship missions should be the first to be put on hold if the planetary-science budget shrinks too much. This would protect the small Discovery missions and medium-sized New Frontiers projects, both of which are relatively low cost, narrowly focused and

selected by competitive peer review. "One message we got from the community was not to let those missions get damaged from overruns on the flagships," says Squyres.

The survey does not prioritize competing Discovery missions — the costs of which are capped at \$500 million — but advocates strongly that the programme should continue to deliver 'economy probes' such as the Mercury surface, space environment, geochemistry and ranging (MESSENGER) mission, which is scheduled to arrive in orbit around the planet next week (see 'MESSENGER's arrival').

The report also suggests adding two projects over the next decade to the New Frontiers programme. These projects cannot exceed \$700 million and include the New Horizons mission, which launched in 2006 to study Pluto and the Kuiper belt. The additions would come from a list of seven candidates, including a sample-return mission to a comet, a Venus lander and an orbiter for Jupiter's moon Io. The committee also strongly endorsed the ESA-NASA Mars Trace Gas Orbiter. Expected to launch in 2016, the orbiter will study the behaviour of trace gases in Mars's atmosphere — particularly methane, which could indicate ongoing geological activity or even the presence of microorganisms.

Although funding helped to dictate the many trade-offs in the survey, Squyres says that the panel relied heavily on input from the planetary-science community and worked to build consensus before making decisions.

Despite some dissension, Squyres urges the planetary-science community to pull together behind the plan, speaking "with one clear and loud voice when talking to elected representatives in Congress to fight for funding".

"We have to be willing to implement some programmes at a reduced cost," says Squyres, "and we have to be willing to fight." ■

TO BOLDLY GO ...

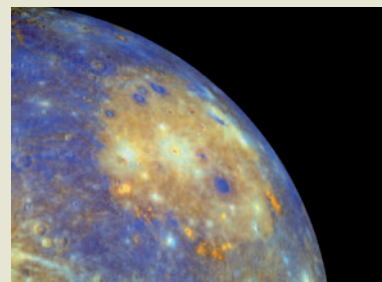
If NASA's planetary-sciences funding remains on track with Obama's 2011 budget request, the decadal survey recommends two flagship missions and a planned joint mission with ESA. It also suggests keeping the Discovery programme intact and choosing two more New Frontiers missions from a shortlist of seven.

Mission	Goals
■ MAX-C (downsized to \$2.5 billion)	Look for evidence of ancient life or prebiotic chemistry; collect samples for return to Earth
■ Uranus Orbiter and Probe	Investigate Uranus, its moons and its rings
■ Mars Trace Gas Orbiter	Study atmospheric trace gases, particularly methane
■ Discovery missions	Cost less than \$500 million, launch roughly every two years
■ Comet Surface Sample Return	Bring substantial comet samples back to Earth
■ Saturn Probe	Investigate Saturn's atmosphere
■ Trojan Tour and Rendezvous	Study multiple Trojan asteroids
■ Venus In Situ Explorer	Land on Venus and analyse environment, including a pristine rock core
■ Lunar South Pole-Aitken Basin Sample Return	Return samples of the Moon's ancient deep crust
■ Io Observer	Understand volcanism on Jupiter's moon Io
■ Lunar Geophysical Network	Study interior of Moon

■ Flagship missions ■ Joint mission ■ Discovery missions ■ New Frontiers missions

DESTINATION MERCURY

MESSENGER's arrival



After a six-year flight, NASA's MESSENGER spacecraft is set to enter orbit around Mercury on 18 March, dipping as close as 200 kilometres to the planet's surface to make a long-awaited survey of this Sun-scorched world.

MESSENGER (Mercury surface, space environment, geochemistry and ranging) is only the second spacecraft to reach the planet. The first, Mariner 10, made three fleeting passes in 1974 and 1975 and glimpsed only 45% of the planet's surface. The new mission has already imaged 98% of Mercury during fly-bys in 2008 and 2009 (see **picture**). Once in orbit, MESSENGER will map the entire surface, recording its topography and composition on scales as small as a few tens of metres. Over the one-year mission, MESSENGER will also study Mercury's tenuous atmosphere of hydrogen, helium and metal ions, and look for water ice in permanently shadowed craters near the poles.

Mercury has an oversized iron core, three-quarters the size of the planet. This is utterly different from other rocky planets, the cores of which are small relative to their mantles. Did Mercury lose some of its mantle after a massive collision with another rocky body, or was it stripped down by an ancient flare-up of the Sun? Each scenario would have left a geochemical fingerprint on the surface, so MESSENGER may settle the matter. A year of magnetometer observations could also reveal whether some of the core is molten.

The wealth of new data expected from MESSENGER isn't likely to be topped for a while, as NASA has no plans for another Mercury mission. BepiColombo, a joint project between the European Space Agency and the Japan Aerospace Exploration Agency, aims to orbit Mercury with two probes simultaneously in 2020, but has cancelled plans for a lander. **Adam Mann**

For more on MESSENGER's mission, see go.nature.com/devj9y

SCIENCE/AAAS, CARNEGIE INST. WASHINGTON/ARIZONA STATE UNIV./JOHNS HOPKINS UNIV. APPL. PHYS. LAB./NASA

ENVIRONMENT

China unveils green targets

Premier vows to improve energy efficiency and curb pollution and carbon emissions.

BY JANE QIU

Growing environmental costs and energy demands have persuaded China's leaders that the country cannot sustain its breakneck economic growth. In a speech on Saturday at the annual National Party Congress in Beijing, Premier Wen Jiabao announced ambitious five-year goals for increasing energy efficiency and curbing carbon emissions — and a reduced target for economic growth.

In the past five years, China's gross domestic product (GDP) has increased at an average rate of 11.2% a year. The country has reduced its energy intensity — or energy consumption per unit of GDP — by 19.1%, just short of a 20% target set five years ago. Emissions of chemical oxygen demand (a measure of organic pollutants in water) and sulphur dioxide per unit of GDP also dipped by 12.5% and 14.3% respectively, exceeding previous targets of 10%.

The latest five-year plan has lowered the economic-growth target to around 7% a year, and calls for energy intensity to decline by a further 16%. The plan is China's first to include targets for carbon emissions per unit of GDP — to be reduced by 17% — and total energy use, which will be capped at 4 billion tonnes of coal equivalent by 2015, compared with the 3.25 billion tonnes consumed last year (see 'Goals for 2015').

"The targets are very ambitious in global terms," says Antony Froggatt, a climate and environment policy expert at Chatham House, a think tank in London that focuses on international affairs. But because the country's imports of coal, oil and gas are increasing each year, he says, the goals "are crucial for China to become competitive on the global market".

The plans are in line with China's existing aim to reduce carbon intensity by 40–45% from 2005 levels by 2020 (see *Nature* 462, 550–551; 2009). Froggatt notes, however, that the commitments fall short of what some climate-policy experts regard as a sufficient contribution towards the current international



China plans to reduce its reliance on coal by upping energy efficiency and the use of non-fossil fuels.

goal of preventing the world's temperature from increasing by 2°C over the pre-industrial level. "But China is not alone in that," he adds.

To meet its targets, China will scrutinize its entire energy-production chain, from suppliers to transmission and end users, and will inspect industries sector by sector, looking for ways to increase efficiency. The country also aims to raise the proportion of its energy coming from non-fossil fuels — notably, wind energy, hydro-power and nuclear power — to 11.4% by 2015. The government hopes that plans to increase spending on research and development to 2.2% of GDP — significantly higher than today — will yield breakthroughs in clean energy.

Critics applaud China's plan to cut energy and carbon intensities, but say that the question is not just whether the targets will be enforced, but how. "Low-carbon economy is not just about number crunching," says Zou Ji, an environmental economist at Renmin University in Beijing. "There are different

FIVE-YEAR PLANNING

Goals for 2015

- Reduce energy and carbon intensity by 16% and 17%, respectively.
- Cap energy use at 4 billion tonnes of coal equivalent.
- Increase the proportion of non-fossil fuels to 11.4% from the current 8.3%.
- Reduce emissions of chemical oxygen demand and sulphur dioxide by 8%.
- Reduce emissions of ammonia and nitrogen oxides by 10%.
- Increase forest coverage to 21.7%.
- Increase timber reserve by 600 million cubic metres.
- Reduce water use per unit of industrial growth by 30%.
- Increase efficiency of irrigation to 53%.

REUTERS

ways to meet the targets, and some are more cost-effective and sustainable than others."

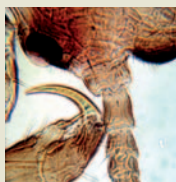
Many say that China met the previous energy-intensity target largely through draconian government measures, such as strict quotas imposed on local governments and energy-intensive industries. Fear of missing their targets led some provinces to severely ration power in the latter half of last year. This "is not necessarily the most cost-effective approach," says Kelly Gallagher, who studies energy and environment policy at Tufts University in Medford, Massachusetts.

The transition to a low-carbon economy requires not only technological innovation, but also drastic reforms in energy pricing, along with mechanisms such as emissions trading and carbon taxation, says Zou. "These are on the agenda of the central government, in principle," he says. "Whether they can be implemented in the next few years will determine the future of China's economic competitiveness." ■



**MORE
ONLINE**

SLIDESHOW



Images from the first microscopes were clearer than was once believed go.nature.com/xvukm6

NEWS

- Satellite monitoring shows oil palms crowding out peat swamps in Malaysia go.nature.com/rmbbrv
- Stretchy electronics could speed up heart surgery go.nature.com/yvusf8
- Is this the end of the Guinness widget? go.nature.com/rkmvr9

ONLINE POLL

How does the Internet influence the way you work? What do you do to manage your online reputation? Take part in *Nature's* online survey about the web and you could win an Amazon gift certificate worth £100/US\$150 go.nature.com/9xpdpn

ENERGY

Cable test raises fears at fusion project

Degradation of superconducting cables for the heart of the ITER fusion machine threatens to cause further delays.

BY GEOFF BRUMFIEL

Scientists on three continents are scrambling to understand a potentially serious problem with superconducting cables destined for ITER, the world's largest fusion experiment. *Nature* has learned that preliminary tests of cable for ITER's powerful central magnet show that it degrades too quickly to be used. If unresolved, the problem threatens to further delay a project that has already suffered many years of technical and budgetary setbacks (see *Nature* **461**, 855; 2009).

ITER is a €15-billion (US\$21-billion) collaboration designed to prove the viability of nuclear fusion as a power source. Now under construction in St-Paul-lez-Durance, France, and scheduled to start power-producing tests in 2026, the reactor will consist of a giant, doughnut-shaped vessel filled with hydrogen isotopes. Powerful magnetic fields generated by huge superconducting coils surrounding the vessel will trap the isotopes as they are heated to the point of fusion.

The largest of these magnets is the central solenoid, a 13.5-metre-high, 4.1-metre-wide stack of six identical coils capable of generating a staggering 13 Tesla magnetic field (see 'ITER's beating heart'). By slowly pulsing the current in the superconducting cable, the solenoid generates a second current inside the hydrogen gas itself, which helps to confine the gas as it heats to fusion temperatures of more than 150 million °C.

The solenoid will be built in the United States using superconducting niobium tin (Nb₃Sn) cable manufactured in Japan. The cable is supposed to withstand 60,000 current pulses without trouble. But tests in November 2010 at the SULTAN facility, part of the Paul Scherrer Institute in Villigen, Switzerland, revealed serious degradation in a sample of cable after only 6,000 pulses. If confirmed, that would mean ITER could operate for only a fraction of its 20-year lifetime with the existing cable.

Mass production of the cable has been put on hold until a solution can be found, says Ryuji Yoshino, the head of Japan's Domestic Agency for ITER. He says that Japanese officials have not seen this kind of degradation in other projects, including earlier ITER test cables produced in 2000. He believes that the problem developed during the Swiss tests, which mimicked ITER's operating conditions by placing the cable in a strong but non-uniform magnetic field.

"I'm concerned, but we have a plan in place to solve the problem," says Ned Sauthoff, the head of the United States' contribution to

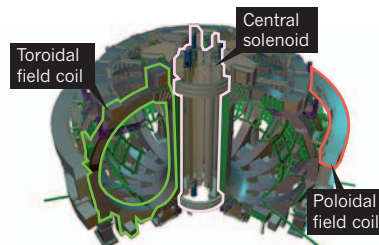
ITER. Scientists at the Japan Proton Accelerator Research Complex in Ibaraki are now conducting neutron-scattering studies to measure the strain on the cable and to check for possible damage. Further experiments will be done in Switzerland to see whether the testing methodology is to blame. Meanwhile, neutron studies of the

failed cable are also under way at the Spallation Neutron Source at Oak Ridge National Laboratory in Tennessee, along with X-ray tomography experiments at Lawrence Livermore National Laboratory in California, according to Sauthoff. In addition, other types of Nb₃Sn cables are being tested to see if they are more resilient. "We're going to test until we get a success," he says.

Sauthoff hopes that the situation can be resolved by June. If a fix for the cable takes longer, then additional facilities could be tasked with manufacturing coils for the central solenoid to speed up production, but this would add extra expense to the project. ■

ITER'S BEATING HEART

The superconducting cables of the central solenoid are a crucial part of the fusion reactor.



CORRECTION

The News story 'Traditional drug-discovery model ripe for reform' (*Nature* **471**, 17–18; 2011) should have put the annual cost of the drug-development model discussed as US\$200 million, not \$325 million.

ENERGY

Cable test raises fears at fusion project

Degradation of superconducting cables for the heart of the ITER fusion machine threatens to cause further delays.

BY GEOFF BRUMFIEL

Scientists on three continents are scrambling to understand a potentially serious problem with superconducting cables destined for ITER, the world's largest fusion experiment. *Nature* has learned that preliminary tests of cable for ITER's powerful central magnet show that it degrades too quickly to be used. If unresolved, the problem threatens to further delay a project that has already suffered many years of technical and budgetary setbacks (see *Nature* **461**, 855; 2009).

ITER is a €15-billion (US\$21-billion) collaboration designed to prove the viability of nuclear fusion as a power source. Now under construction in St-Paul-lez-Durance, France, and scheduled to start power-producing tests in 2026, the reactor will consist of a giant, doughnut-shaped vessel filled with hydrogen isotopes. Powerful magnetic fields generated by huge superconducting coils surrounding the vessel will trap the isotopes as they are heated to the point of fusion.

The largest of these magnets is the central solenoid, a 13.5-metre-high, 4.1-metre-wide stack of six identical coils capable of generating a staggering 13 Tesla magnetic field (see 'ITER's beating heart'). By slowly pulsing the current in the superconducting cable, the solenoid generates a second current inside the hydrogen gas itself, which helps to confine the gas as it heats to fusion temperatures of more than 150 million °C.

The solenoid will be built in the United States using superconducting niobium tin (Nb₃Sn) cable manufactured in Japan. The cable is supposed to withstand 60,000 current pulses without trouble. But tests in November 2010 at the SULTAN facility, part of the Paul Scherrer Institute in Villigen, Switzerland, revealed serious degradation in a sample of cable after only 6,000 pulses. If confirmed, that would mean ITER could operate for only a fraction of its 20-year lifetime with the existing cable.

Mass production of the cable has been put on hold until a solution can be found, says Ryuji Yoshino, the head of Japan's Domestic Agency for ITER. He says that Japanese officials have not seen this kind of degradation in other projects, including earlier ITER test cables produced in 2000. He believes that the problem developed during the Swiss tests, which mimicked ITER's operating conditions by placing the cable in a strong but non-uniform magnetic field.

"I'm concerned, but we have a plan in place to solve the problem," says Ned Sauthoff, the head of the United States' contribution to

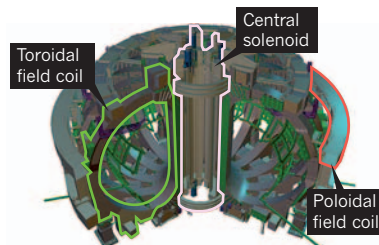
ITER. Scientists at the Japan Proton Accelerator Research Complex in Ibaraki are now conducting neutron-scattering studies to measure the strain on the cable and to check for possible damage. Further experiments will be done in Switzerland to see whether the testing methodology is to blame. Meanwhile, neutron studies of the

failed cable are also under way at the Spallation Neutron Source at Oak Ridge National Laboratory in Tennessee, along with X-ray tomography experiments at Lawrence Livermore National Laboratory in California, according to Sauthoff. In addition, other types of Nb₃Sn cables are being tested to see if they are more resilient. "We're going to test until we get a success," he says.

Sauthoff hopes that the situation can be resolved by June. If a fix for the cable takes longer, then additional facilities could be tasked with manufacturing coils for the central solenoid to speed up production, but this would add extra expense to the project. ■

ITER'S BEATING HEART

The superconducting cables of the central solenoid are a crucial part of the fusion reactor.



CORRECTION

The News story 'Traditional drug-discovery model ripe for reform' (*Nature* **471**, 17–18; 2011) should have put the annual cost of the drug-development model discussed as US\$200 million, not \$325 million.

BREAKING THE PROTEIN RULES

If dogma dictates that proteins need a structure to function, then why do so many of them live in a state of disorder?

BY TANGUY CHOUARD

Keith Dunker's life is a mess. His desk is so swamped with books, old chocolate bars, half-reviewed manuscripts, pens, diet coke bottles and — somewhere — a stray sock, that he ends up printing papers again rather than wading in to find the original. "I'm so disorganized," he crows, "some people have called me Dr Disorder." But he remembers with great precision the moment that disorder invaded his scientific life. It was 15 November 1995, at 12:40 p.m., halfway through a seminar by crystallographer Chuck Kissinger, at Washington State University in Pullman, where Dunker was then a biochemist. Dunker was staring at a slide showing the atomic structure of calcineurin, an enzyme targeted by immunosuppressive drugs. What caught his attention wasn't the intricate structure but something missing from it: a dotted line representing a string of amino acids with a position too variable to be determined by X-ray crystallography, as the rest of the protein had been. And Kissinger was insisting that this loop had to remain flexible for calcineurin to serve its crucial function in the human immune system.

"It hit me like a brick," says Dunker: this wayward piece of protein flouted a century of dogma. A central tenet in molecular biology is that the function of a protein depends critically on its fixed three-dimensional structure; for example, enzymes bind to specific substrates because their shapes match perfectly, as immortalized in the 'lock-and-key' model proposed by chemist Emil Fischer as early as 1894. But this part of calcineurin seemed to disobey these rules, by providing function without structure. Now Dunker was wondering how many other proteins were ignoring the rules too.

To find out, he and his colleagues wrote a bioinformatics program that predicted which protein segments are 'intrinsically disordered' — meaning that they do not fold spontaneously into a unique three-dimensional shape. Today, this and other similar programs predict that about 40% of all human proteins contain at least one intrinsically disordered segment of 30 amino acids or more, and that some 25% are likely to be

disordered from beginning to end¹. This part of the protein universe had largely been ignored because disordered protein segments impede crystal formation — a prerequisite for X-ray diffraction, the predominant way structures are deduced — and structural biologists clip them out whenever they can.

Today, though, "the recognition of disorder has grown dramatically", Peter Wright, a protein biophysicist at the Scripps Research Institute in La Jolla, California, told the American Association for the Advancement of Science meeting in Washington DC last month. A large part of that recognition has come from studies using nuclear magnetic resonance (NMR) spectroscopy, which allows researchers to determine the structures of small proteins even as they twist and turn in solution. Such work has shown that disorder can actually be essential to function by helping a signalling protein to recognize and react to a protein partner, or by allowing a regulatory protein to interact with multiple targets. Still, says Wright, "that hasn't got through to the textbooks".

Many structural biologists see no need for revision. "My mantra has been: function requires structure," declares Tom Steitz, a crystallographer at Yale University in New Haven, Connecticut. "Some flexibility can be required, it may be an essential part of the assembly process, but it's not interesting until the proteins get to do their job." Critics argue that the computer programs predicting high levels of disorder are fundamentally flawed because they identify proteins that are well-known to become perfectly ordered — and to crystallize — when bound to their proper molecular partners. They say that unfolded protein chains cannot persist for long in living cells, and some want the concept of intrinsically disordered proteins to be ditched altogether. That seems unlikely. Data are fast accumulating from all fronts — biophysics, bioinformatics and cell biology — in support of widespread disorder, and disorder aficionados are calling for a complete reassessment of the structure–function paradigm. "Biology uses disorder to bring about its various functions," Wright says.

Since the late 1950s, newly made proteins have been assumed to fold up immediately and spontaneously into a unique three-dimensional shape — their most energetically stable conformation and the only functional

➔ **NATURE.COM**
Read about protein
folding that can
cause disease:
go.nature.com/s6w5mm

one². The few proteins known to remain unfolded “were pointed to as oddities”, Wright says. But that started to change in 1999, when Wright and fellow NMR spectroscopist Jane Dyson, also at Scripps, wrote a review³ pointing to the growing collection of proteins that seemed to function despite their disordered state. It has been “the big-dog paper in the field”, says Dunker, now at Indiana University School of Medicine in Indianapolis.

One burning question, then and now, is how a protein can function if it has no fixed shape. “We all accept flexibility,” says structural biologist Joël Janin at the CNRS Laboratory of Enzymology and Structural Biochemistry at Gif-sur-Yvette, France. “The question is: how can you get recognition with flexibility?” The whole concept of disorder seems incompatible with the lock-and-key model. You might as well try to open the door with cooked spaghetti.

In 2007, postdoc Kenji Sugase in Wright’s lab found an answer: the spaghetti uses the lock to mould itself into the shape of the key, rather than forming the key beforehand. Sugase focused on CREB, a gene-regulatory protein involved in many processes including learning and memory. Once bound to DNA, CREB also needs to recognize and bind a protein partner called CBP before it can switch on a gene. But the part

of CREB involved with CBP enters the game in a disordered state. How could a thing like this possibly work?

To find out, Sugase developed the equivalent of a super-fast NMR camera so that he could capture frequent snapshots of CREB’s wriggling chain, at atomic resolution, as it tested out points of contact within itself and with CBP. What he saw was that several bonds had to form cooperatively within CREB and with CBP for the whole complex to snap into shape⁴. That’s exactly how the average ‘globular’ protein folds: its internal segments need to establish long-range chemical bonds with one another, to pull the whole thing suddenly into shape². CREB forms such interactions externally, by bonding to CBP — and if this bonding is just a little weaker, then the key cannot form and there is no binding with the lock at all. Wright and his colleagues think that disorder is therefore advantageous because it allows CREB to partner with CBP more exclusively than a rigid protein would. And Wright thinks that this type of process allows many signalling proteins to engage in fast-changing yet selective interactions.

DISORDERLY CONDUCT

The structure–function mantra took an even bigger hit recently from a protein with parts that never seem to fold at all. The Sic1 signalling protein is a key regulator of the cell cycle that puts the brakes on DNA replication until the cell is ready to divide. In 2001, a team led by Mike Tyers, a yeast cell-cycle expert at the University of Toronto, Canada, began unpicking the mechanism of the switch. The group found that when phosphate groups are added to six sites on Sic1 it can then hook up with a second protein, Cdc4, which pushes Sic1 into the cell’s protein-disposal pathways⁵. Once Sic1 is degraded, DNA replication can forge ahead. But unless that degradation occurs at precisely the right time, DNA replication goes haywire and the cell may eventually die. The cell achieves that precision by ensuring that it takes exactly six phosphates to flip the switch, not four or five. But there’s a rub: Cdc4 has only one high-affinity binding pocket for a phosphate group. How can Cdc4 ‘count’ up to six with effectively only one finger to count on?

After all attempts to crystallize Sic1 failed, Tyers’s team called NMR spectroscopist Julie Forman-Kay, also at the University of Toronto. In 2008, Tanja Mittag, a postdoc in Forman-Kay’s lab, showed that Sic1 was disordered⁶ — not only in its free state but, astoundingly, also when bound to Cdc4. The complex seemed to be a mixture of different conformations shifting around in constant, dynamic equilibrium. And the most stunning part was that each of the six phosphate groups on Sic1 could be found to occupy the single Cdc4 pocket, one after the other, as in a constant dance around the fire (see ‘Orders of disorder’).

The researchers then developed a computer model and fed it with every scrap of experimental data about the proteins’ structures that they could gather⁷. They concluded that, even though Sic1 is disordered when it is bound, it maintains a rather compact structure, which keeps all the phosphate groups sufficiently close together to form an average electrostatic field that glues Sic1 to Cdc4. Only when six phosphates are present is the glue strong enough for Cdc4 to hold Sic1 close and force-feed it into the cell’s disposal machinery. And one reason that Sic1 has to be disordered during all this, the team proposes, is to enable the rather rigid disposal machinery to reach all parts of Sic1 and carpet it with the chemical tags that mark the protein for destruction. It takes a nimble protein indeed to make all these connections at once.

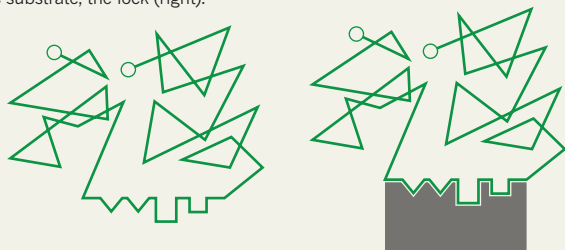
“The result is interesting,” says structural biologist Stephen Harrison of Harvard Medical School in Boston, “because interaction motifs are often found more or less repeated along unstructured segments, and the work shows how such multiplicity can function.” And Forman-Kay thinks that adopting ‘multi-structural’ states could allow other proteins to constantly probe and sense signals from many partners at once. This is particularly important for ‘hub’ proteins, which are central to vast networks of rapidly changing molecular interactions. “There is a complexity people haven’t talked about,” she says. “These hub proteins need to very rapidly sample the complex cellular environment.”

ORDERS OF DISORDER

A central tenet in molecular biology states that the function of a protein depends on its fully folded three-dimensional structure. In the new view, protein segments can function when transiently or durably disordered.

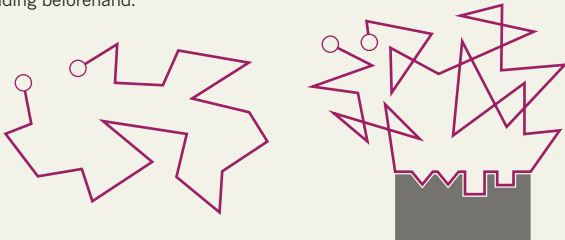
LOCK AND KEY

In the conventional view, an enzyme folds up immediately into a unique and stable 3D shape, the key (left). Its shape perfectly matches and allows it to bind its substrate, the lock (right).



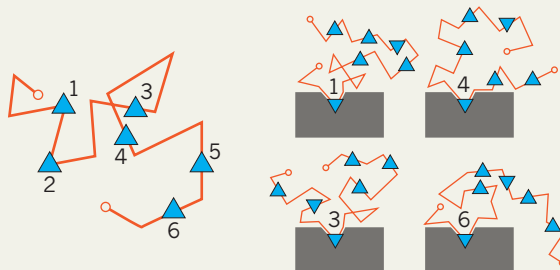
FOLD AS YOU BIND

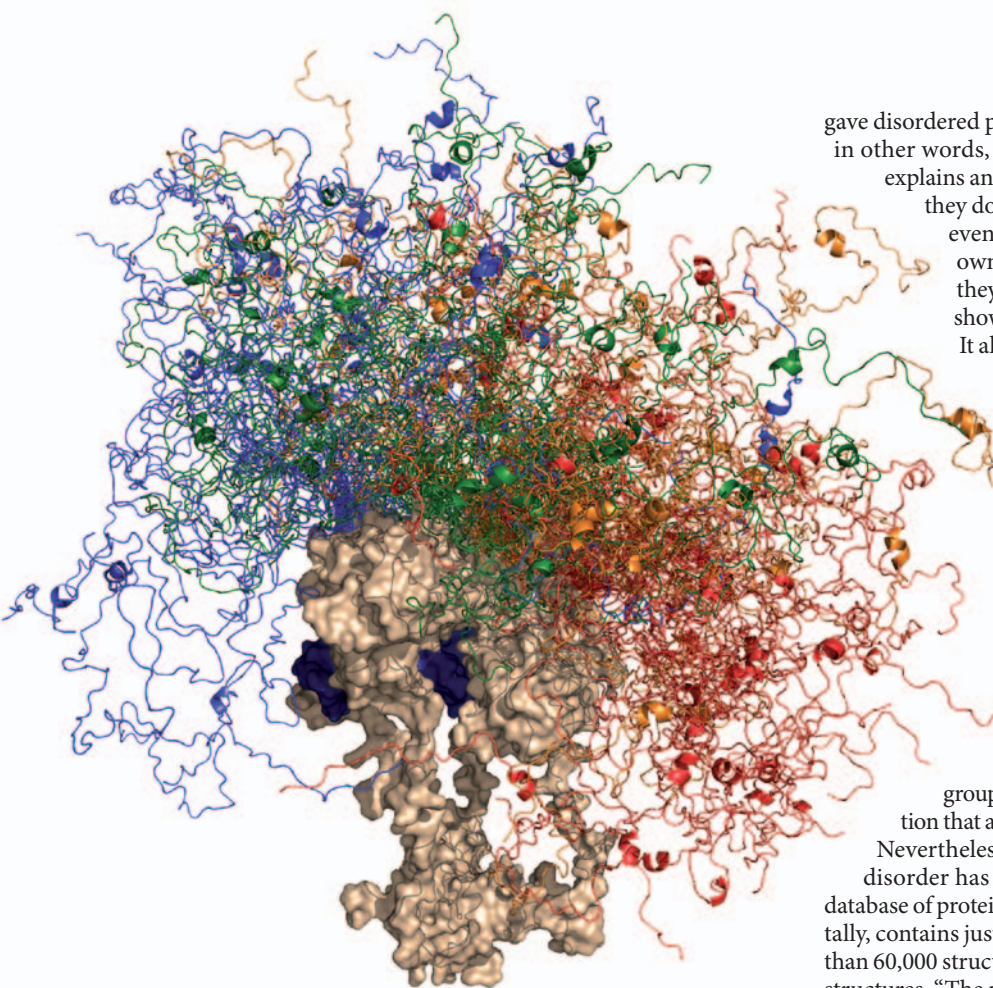
A disordered part of the gene-regulatory protein CREB (left) uses the lock to mould itself into the shape of the key when the two meet (right), rather than folding beforehand.



SHAPE SHIFTING

The signalling protein Sic1 remains disordered in its bound state, and each of six phosphate groups occupies the binding site in turn. The protein is a mix of different conformations shifting around in constant dynamic equilibrium.





Tumour-suppressor protein p53 contains an ordered, globular domain (brown); its disordered segments (colours) help it interact with hundreds of partners.

One extreme example can be seen in the tumour suppressor p53, an extraordinarily well connected hub in multiple signalling networks, and the protein most frequently implicated in human cancer. Part of the explanation for p53's promiscuity seems to lie in its versatile structure, which features every possible conformation from order to disorder. The core domain is globular and binds to DNA and just a few other proteins; its two flanking wings are mostly disordered and can bind to hundreds of signalling partners; and a segment within one wing shows a 'chameleon' status which can flip between four different ordered states, depending on which partner it binds to⁸. Alan Fersht, a biophysicist and NMR expert at the University of Cambridge, UK, says that he is "absolutely sure" that long parts of p53 remain largely disordered in the cell: "I don't think there's any doubt about that whatsoever."

PREDICTION PROBLEMS

Yet many researchers question how widespread disordered proteins can be. That is mainly because biochemists, over more than 100 years of preparing tissue extracts, have struggled to prevent proteins from unfolding and tangling into insoluble clumps or getting digested by enzymes called proteases. "It's hard to believe that disordered proteins could produce anything else than a mess," says Janin.

Researchers who study these processes, however, say that such fears are largely irrelevant. In human cells, for example, nonspecific proteases are locked away in compartments called lysosomes, which should allow disordered proteins to survive everywhere else, explains Ulrich Hartl, an expert on protein quality control at the Max Planck Institute in Martinsried, Germany. Disordered proteins should also be protected from aggregation because, unlike globular proteins, they contain few hydrophobic amino acids, which tend to stick together — and are instead rich in 'polar' amino acids that are happy swimming in water.

Hartl thinks that natural selection against aggregation probably

gave disordered proteins this particular amino-acid composition — in other words, it is not a signature for disorder per se. And this explains an apparent inconsistency of disorder predictors: that they do not pick internal segments from globular proteins, even though these too are incapable of folding on their own when sliced out of a protein². The reason is that they are rich in hydrophobic amino acids and so do not show the signature sequence that the predictors detect.

It also explains why the predictors select some proteins that are in fact ordered, another point of controversy surrounding these programs: because proteins can lack hydrophobic amino acids for many different reasons. "This makes perfectly good sense," agrees Dunker.

Overall, the numerous programs claim an 80% success rate at predicting whether any individual amino acid in a protein will be surrounded by order or disorder¹, as compared to the 50% success rate expected by chance — and crystallographers rely on them to sideline proteins expected to resist crystallization. "Disorder predictors are a massive oversimplification, but they are very useful," says Adam Godzik, at the Burnham Institute in La Jolla, California, and head of the bioinformatics group for the Protein Structure Initiative, a large collaboration that aims to solve large numbers of protein structures.

Nevertheless, debate about the prevalence and importance of disorder has probably slowed progress in the field. DisProt, a database of proteins whose disorder has been established experimentally, contains just over 500 proteins, a number dwarfed by the more than 60,000 structures in the Protein Data Bank¹, the database for 3D structures. "The main reason why DisProt is so small," says Dunker, "is because it has taken so many damn years to get it funded." But in the past few years, major consortia aimed at exploring intrinsically disordered proteins have been set up in several countries. Interest in disordered proteins as drug targets is also on the rise because so many of them, like p53, are crucially implicated in disease.

Little by little, a fundamentally new picture of the relationships between protein sequence, structure and function is emerging: a continuum running from the most rigid 'lock-and-key' enzymes and molecular machines at one extreme through to durably unstructured spaghetti such as Sic1 at the other, and spanning all degrees of structural ambiguity in between. Figuring out how all these disordered proteins really work is a long way off, if Sic1 is anything to go by: determining its mode of action involved several, often arcane, biophysical techniques, new computer tools and statistical physics theory — plus at least ten years of work by six labs. Multi-structural biology isn't going to be simple.

Still, Dunker, Wright and other doctors of disorder are optimistic. As is Martin Blackledge, an NMR spectroscopist at the Institute of Structural Biology in Grenoble, France, who compares the excitement now to that surrounding the first crystal protein structures in the 1950s. "Every new case is fascinating at the moment," he says. Blackledge looks forward to a day when it might be possible to predict where on the structural continuum a protein segment falls, from its amino-acid sequence — to crack the full code of disorder. "This is exactly what I'm aiming for," he says, "this is my dream."

Perhaps the rules of disorder are needed, before disorder can rule. ■

Tanguy Chouard is an editor for *Nature* in London.

1. Uversky, V. N. & Dunker, A. K. *Biochim. Biophys. Acta* **1804**, 1231–1264 (2010).
2. Anfinsen, C. B. *Science* **181**, 223–230 (1973).
3. Wright, P. E. & Dyson, H. J. *J. Mol. Biol.* **293**, 321–331 (1999).
4. Sugase, K., Dyson, H. J. & Wright, P. E. *Nature* **447**, 1021–1025 (2007).
5. Nash, P. et al. *Nature* **414**, 514–521 (2001).
6. Mittag, T. et al. *Proc. Natl Acad. Sci. USA* **105**, 17772–17777 (2008).
7. Mittag, T. et al. *Structure* **18**, 494–506 (2010).
8. Oldfield, C. J. et al. *BMC Genomics* **9** (Suppl. 1), S1 (2008).



EARTH'S ACID TEST

BY QUIRIN SCHIERMEIER

As the oceans rapidly grow more acidic, scientists are scrambling to discover how marine life is likely to react.

The Friday night beers made Sam Dupont forget all about his sea urchins. Earlier that day, in April 2010, the young Belgian eco-physiologist had put a batch of urchin larvae into a bath of highly acidic water to see how their skeletons would fare.

J.-P. GATTUSO/CNRS

When nothing obvious happened after a few hours, Dupont decided to join some friends at the pub and check on the experiment later in the evening. But he didn't remember until Sunday, at which point he was sure that the precious larvae would be dead.

But when Dupont returned to work at the Sven Lovén Centre for Marine Sciences in Kristineberg, Sweden, on Monday, he found the larvae still swimming around in their tank. Their internal skeletons had dissolved away, but otherwise the creatures seemed to be functioning well.

Dupont's chance finding underscores how much scientists have yet to learn about the growing threat of ocean acidification, which is caused by rapidly rising atmospheric concentrations of carbon dioxide. The acidity of sea water has climbed by 30% over the past 150 years, and some regions have already become corrosive enough to inhibit the growth of corals and other species for part of the year. According to projections, most creatures with calcium carbonate shells, such as mussels and snails, could run into problems within a few decades. By the end of this century, the acidification could even impede the growth of important groups of plankton, thus endangering entire marine ecosystems, from fisheries to coral reefs.

Although the urchin experiment hints that some organisms are able to survive brief exposures to highly acidic water, other studies are revealing unexpected problems that might threaten even creatures without hard shells, such as fin fish. Preliminary work suggests that responses could be highly variable, depending on factors such as water temperature, a creature's evolutionary history and the availability and quality of food.

An experiment off the coast of Spitsbergen tests the effects of elevated carbon dioxide concentrations on marine life.

Countries are only now revving up the coordinated research programmes needed to assess how marine ecosystems will react to the increasingly acidic waters. “We simply have not conducted the basic experiments,” says Richard Feely, an oceanographer with the US National Oceanic and Atmospheric Administration in Seattle, Washington, which last year launched a US\$5.5-million programme of research into the problem. But with the current pace of acidification, scientists do not have much time to come up with answers.

CARBON SINK

Without the oceans and their vast ability to absorb carbon dioxide, Earth would be warming up much faster than it currently is. The seas take up about 9 billion tonnes of the gas each year — almost one-third of the 30 billion tonnes emitted globally.

Once it enters the ocean, CO_2 reacts with water to produce carbonic acid, which releases positively charged hydrogen ions. Acidity is measured in pH, a logarithmic scale on which low numbers mean high acidity; neutral water has a pH of 7, but sea water is naturally alkaline, owing to the salts dissolved in it. Since the mid-nineteenth century, the average pH of ocean surface waters has dropped by 0.1 units, to a current value of about 8.1. Unless nations sharply curb their emissions, atmospheric CO_2 is expected to at least double from its preindustrial concentration by sometime in the second half of this century, and scientists project that ocean pH will fall by a further 0.3–0.4 or so units. Sea water could then contain at least 150% more hydrogen ions than it did at the onset of the industrial era.

Those extra ions cause problems by binding with dissolved carbonate ions to form bicarbonate. With fewer free carbonate ions in the water, organisms struggle to absorb enough to build shells and skeletons made of calcite and aragonite — two different forms of calcium carbonate. And if sea water becomes permanently undersaturated with respect to those minerals, hard parts made of them will start to dissolve.

“There is absolutely no doubt that calcifying organisms will calcify less if conditions become more acidic,” says Jean-Pierre Gattuso, an oceanographer at the National Centre for Scientific Research in Villefranche-sur-mer, France, who coordinates the European Project on Ocean Acidification (EPOCA).

This has happened before. Some 55 million years ago, during an episode of extreme global warming driven by a spike in atmospheric CO_2 , the pH of sea water is thought to have dropped to levels similar to those expected at the end of the twenty-first century. Ocean sediment deposited during that period contains very little carbonate and no fossils of microorganisms with calcium carbonate shells, indicating that the sea water became too corrosive for calcifying algae such as deep-sea foraminifera, driving many to extinction¹. Today, acidification is progressing at least ten times faster than it did 55 million years ago.

Researchers expect to see problems pop up first in polar seas, because cold water absorbs more CO_2 than warmer water (and because the melting of sea ice dilutes the concentration of carbonate ions). In 2008, measurements showed that regions of the Arctic Ocean had become undersaturated with respect to aragonite for part of the year², and scientists suggest that further portions of the Arctic and Southern Oceans will cross that chemical threshold within the next decade. If CO_2 continues to rise at current rates, half of the Arctic Ocean could be undersaturated with respect to aragonite year-round by 2050 (see ‘Into the red zone’).

Even in temperate waters, pH changes may already be having an impact. In the United States, the West Coast shellfish industry has asked scientists to study a dramatic rise in oyster mortality seen in hatcheries off Oregon and Washington since 2005.

During the summer, upwelling currents in these seas carry deep-ocean water, naturally under-saturated with respect to calcium carbonate, onto the continental shelf. Researchers wonder whether the acidification of surface waters has combined with these upwelling currents to cause some of the recent shellfish problems.

At the moment, scientists can offer few conclusions. Although they can make broad predictions about the progress of ocean acidification, they know very little about how it will affect marine animals in different climate zones, alter the composition of ecosystems and, ultimately, influence the marine food web.

To complicate matters, acidification is just one of many environmental changes confronting marine life. Organisms also face increasing stress from ocean warming, pollution, fishing pressure, sea-ice loss and shifting patterns of currents and mixing of deep and shallow water. Some scientists think that progressive ocean acidification will limit the ability of marine organisms to survive such stresses.

SEA OF VARIABLES

Dupont and his colleagues in Kristineberg tried to answer some of the basic questions about acidification by filling 264 tanks with a range of organisms, including scallops, halibut, brittle stars, sea urchins and lobsters. In a four-month lab experiment that ended this week, they observed the performance of the various animals in each combination of six temperatures (6–18 °C) and two pH values (8.1 and 7.7), measuring growth, respiration, shell and tissue structure, internal pH and survival rates. They are just starting to analyse the data.

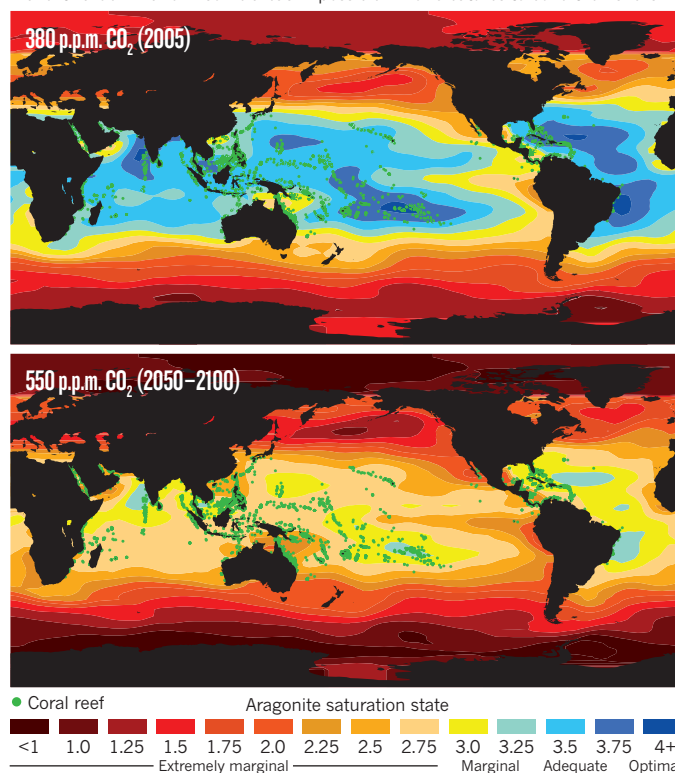
With his previous urchin test, Dupont says, “we’ve seen that some species can cope with extremely low pH values, at least in the short term”. But that might not be true for longer exposures and higher temperatures. “We expect that the response to combined stressors is very site- and species-specific.”

A separate study of two populations of spider crabs (*Hyas araneus*) suggests that how animals respond to acidification depends on their climate zone. In lab experiments, the growth rate and fitness of larvae from the North Sea decreased markedly in acidic waters, whereas an Arctic population from 3,000 kilometres farther north was more sensitive to warming than increased acidity³.

Even individuals from the same species and climate zone can react

INTO THE RED ZONE

Models suggest that rising levels of atmospheric carbon dioxide will reduce the oceans’ saturation state of aragonite, a mineral in the shells of marine organisms. That will make it harder — and in some areas impossible — for creatures to build their shells.

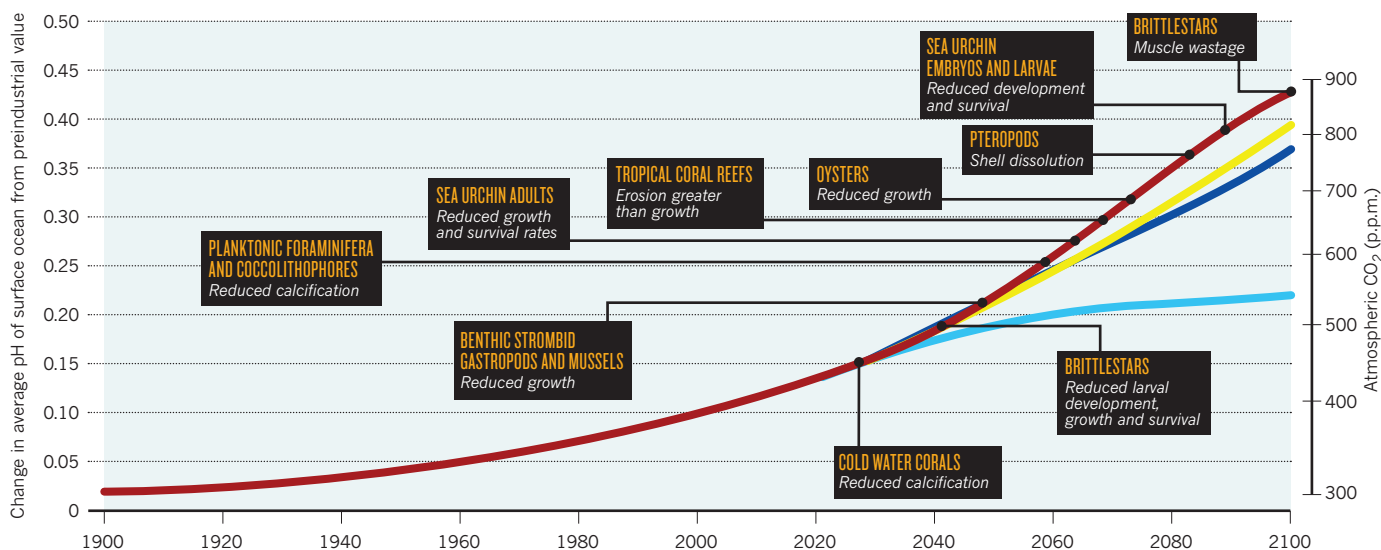


SOURCE: L. CAO & K. CALDERA GEOPHYS. RES. LETT. 35, 119609 (2008)/REEFS AT RISK REVISITED (WRI, 2011)

NATURE.COM
For more on Earth's
environmental
boundaries, visit:
go.nature.com/ejfwxi

FUTURE SHOCKS

Experiments suggest that marine organisms will respond differently to rising ocean acidity, depending on their physiology and habitat. The speed of acidification, and the timing of effects, will depend on future emissions of carbon dioxide (four scenarios are shown, in different colours).



SOURCE: C. TURLEY ET AL. MAR. POLLUT. BULL. 60, 787–792 (2010)

quite differently. In one lab study⁴, blue mussels (*Mytilus edulis*) from the North Sea showed a 25% drop in calcification rates at values of atmospheric CO₂ of 740 parts per million, about what is expected by 2100 if emissions are not curbed. But a different population seemed to do just fine in such waters: these mussels live in the nearby Baltic Sea, where CO₂-rich waters well up for parts of the year, causing the pH in the sea to drop as low as 7.5 (ref. 5). Frank Melzner, an environmental physiologist at the Leibniz Institute of Marine Sciences (IFM-GEOMAR) in Kiel, Germany, who led the Baltic study, suggests that the mussels can survive there because they have developed the physiological capacity to regulate the pH in their cells and build up a protective layer of proteins and carbohydrates that shelters their shells. But only well-nourished organisms can afford such defences, he says. “It seems that some organisms can biologically control the effects surprisingly well — but it certainly requires energy.”

There is plenty of food in the Baltic. Where nutrition is less abundant, populations seem to decline when faced with increased acidity. That is one of the lessons from a study off the Italian island of Ischia in the Gulf of Naples, where underwater volcanic vents have been spewing CO₂ into the comparatively food-poor Tyrrhenian Sea for millennia. A survey of life around the site found that normally common calcifying organisms, including corals and sea urchins, were absent from the spots with low pH. Instead, the researchers discovered a thriving community of species that are immune to elevated CO₂ or even benefit from it, such as sea grasses and invasive algae⁶.

Jason Hall-Spencer, a marine biologist at the University of Plymouth, UK, who oversees the research off Ischia, says that the massive difference between the responses of animals there and in the Baltic illustrates how little is known. To really understand the problem, “you’d like to test the combined effects of ocean acidification and other stressors on hundreds of species and their interactions”, he says.

And calcifying organisms are not the only creatures at risk (see ‘Future shocks’). Even fish could be vulnerable: experiments have shown that elevated CO₂ impairs the sense of smell in juvenile clownfish (*Amphiprion percula*), which could make it difficult for them to find the sea anemones in which they like to live⁷.

GROWING URGENCY

In the past few years, nations have started to devote resources to the research challenge. Europe’s €16.5-million (US\$22.9-million), four-year EPOCA project, which began in 2008 and encompasses 31 laboratories in 10 countries, aims to monitor the effects of ocean acidification on

marine organisms at various scales, from cells to ecosystems and then across the entire globe. One of the programme’s priorities is to determine whether there are any tipping points, beyond which any increase in acidity would hurl marine ecosystems towards catastrophic changes.

In the United States, President Barack Obama’s administration plans to submit a proposal to Congress in the next month or so for an integrated national research programme on ocean acidification, which would draw together researchers from across the federal government. The president’s 2011 budget called for \$11.6 million for research on the subject, but Congress has yet to pass a budget for the current fiscal year. National research programmes are also under way in Germany, Britain, Japan, China, South Korea and Australia.

The largest field experiment conducted so far is an offshore study by EPOCA, involving algae and bacteria in large floating containers exposed to varying levels of CO₂. The research took place between May and July last year, off the island of Spitsbergen in the Arctic Ocean. A group of 35 researchers collected daily measurements of 45 variables affecting the ‘mesocosm’ within the oversized containers, from nutrient cycling to trace-gas production by calcifying algae. The experiment is to be repeated in April and May this year off Bergen in Norway, where the team hopes to observe how acidification affects a bloom of coccolithophorids — important calcifying algae that produce dimethyl sulphide, a trace gas that seeds the formation of clouds.

“We need to understand much more about how ocean acidification affects real ecosystems than we can hope to learn from dose-response experiments on isolated species,” says Ulf Riebesell, a biological oceanographer at the IFM-GEOMAR, who leads the study.

A sense of urgency is propelling these studies. Governments have shown no signs of stemming CO₂ emissions any time soon, and there is talk of tackling the problem of methane and other greenhouse gases first, leaving the tougher issue of CO₂ for a later generation. That might slow the global temperature rise, but it won’t keep the seas from growing ever more corrosive. ■

Quirin Schiermeier is a senior reporter with Nature in Munich.

1. Zachos, J. C. et al. *Science* **308**, 1611–1615 (2005).
2. Yamamoto-Kawai, M., McLaughlin, F. A., Carmack, E. C., Nishino, S. & Shimada, K. *Science* **326**, 1098–1100 (2009).
3. Walther, K., Anger, K. & Pörtner, H. O. *Mar. Ecol. Progr. Ser.* **417**, 159–170 (2010).
4. Gazeau, F. et al. *Geophys. Res. Lett.* **34**, L07603 (2007).
5. Thomsen, J. et al. *Biogeosciences* **7**, 3879–3891 (2010).
6. Hall-Spencer, J. M. et al. *Nature* **454**, 96–99 (2008).
7. Munday, P. L. et al. *Proc. Natl Acad. Sci. USA* **106**, 1848–1852 (2009).

COMMENT

COMMUNICATION Publicize biobanks or they could fail, surveys suggest **p.159**

BIOGRAPHY A remarkable journey through post-war physics **p.161**



HISTORY Magisterial series on Islamic science concludes **p.162**

ECONOMICS Why such complex models when practical solutions are to hand? **p.166**



J. LAMPEN/REUTERS/CORBIS

Vaccinate for the next H2N2 pandemic now

An old influenza strain still circulating in birds and swine could easily jump back to humans now that immunity to it has dropped, warn **Gary J. Nabel** and his colleagues.

The emergence of a new strain of H1N1 influenza virus in 2009 took the world by surprise. The public-health community had assumed that a pandemic strain would arise from a major genetic reshuffling if RNA from a seasonal strain recombined with RNA from a virus that had never circulated in humans before.

As it turned out, the virus bore a remarkable resemblance to one that had already caused a pandemic — 90 years earlier. The major surface protein from the 2009 H1N1 was strikingly similar to the same type of protein from the 1918 H1N1 Spanish flu virus^{1–4}, which killed about 50 million people

worldwide. Indeed, last year, researchers discovered that antibodies able to prevent the 1918 strain from entering cells in mice had the same effect on the 2009 H1N1^{1–3}.

Although the 1918 virus has long since evolved into widely divergent seasonal strains, a version with a very similar surface protein has circulated in pigs for nearly a century. It was therefore poised to cross back into humans and cause a new pandemic when broad protective human immunity had waned.

This unexpected source of the 2009 H1N1 pandemic is a cautionary tale for the public-health community. Another subtype of

influenza, H2N2, looms as a public health threat^{5–7}, and could re-emerge in a similar way. Governments, regulatory agencies and industry should develop a pre-emptive vaccination programme for H2N2.

Like the 1918 virus, H2N2 influenza has already demonstrated its ability to cause a pandemic. From 1957 until 1968, an H2N2 strain caused between 1 million and 4 million deaths worldwide. Also like the 1918 strain,

H2N2 viruses have not circulated in humans for several decades, but continue to do so among birds and swine⁸. The ►

► **NATURE.COM**

See *Nature's* swine-flu special: nature.com/swineflu

▶ oldest human H2N2 strains are closely related to avian strains, suggesting that the 1957 pandemic arose from birds. In fact, the H2N2 subtype mutates relatively slowly: in most bird and human strains, the surface proteins are 92% or more identical.

To examine levels of immunity to this class of virus, between 2003 and 2007, we tested for antibodies against H2N2 strains in the blood of a small cohort of 90 people in the United States. Ideally, this assay should be repeated in several thousand individuals, but our study suggests that people under the age of 50 have little or no immunity, and resistance dramatically increases for those older than 50 (see 'Vulnerability of youth'). This was also the case for the 2009 H1N1.

The low mutation rate for H2N2, and evidence of waning human immunity, make it likely that an H2N2 pandemic could arise from animals⁵⁻⁷. What steps can be taken to prevent its re-emergence?

EXISTING VACCINE

The genetic similarity of the circulating H2N2 strains and the fact that antibodies effective for one strain can work for others suggest that vaccines previously used for an H2N2 virus are likely to protect people against future pandemics^{5,6,9}. Indeed, the vaccine licensed in 1957, and administered until the late 1960s, can protect mice against currently circulating animal H2N2 strains^{5,9}.

However, the manufacturing process for vaccines is unpredictable and costly, making it difficult to conduct a reactive but timely response. For example, by the time vaccines against the 2009 H1N1 virus were made available in most industrialized countries, the outbreak was past its peak. The result of this delay was an estimated 1 million extra infections, stressed health systems and tens of millions of unused vaccine doses¹⁰.

For H2N2 — for which a safe vaccine that is likely to be effective is already available — a pre-emptive vaccination strategy¹⁰ is far preferable, in our view. There are several ways to proceed.

One approach would be to manufacture the vaccine licensed in 1957 and immunize enough of the world's population to provide 'herd immunity' to the rest. This could be achieved by a 'one-time' campaign to immunize most of the adult population worldwide — for example, as part of standard seasonal flu vaccinations — accompanied by an ongoing programme to administer the vaccine to children. Currently, immunizing

10 million people in the United States costs just US\$250 million or less.

Another approach is to stockpile the vaccine so supplies are ready in the event of an outbreak. This would allow health practitioners to deploy the vaccine faster than they could if there was no preemptive strategy in place. This is likely to be more expensive and less effective than routine vaccination. The inevitable delay in distributing the vaccine would allow the virus to spread, kill more people and potentially mutate to the point of being able to evade people's immune systems. Maintaining large amounts of quality-controlled vaccine could also cost tens to hundreds of millions of dollars each year in the United States alone, because its limited shelf life means that it would need to be replenished.

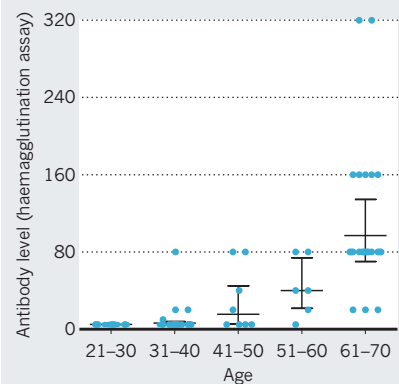
A third possibility is to make 'master lots' of the H2N2 vaccine and ramp up production as soon as signs of an outbreak occur. Although this approach may be cheaper, it is less likely to be effective than either of the above. The 2009 epidemic demonstrated that it takes time to distinguish a serious outbreak from ordinary seasonal fluctuations and to identify the agent responsible.

Some argue that it is impossible to justify politically the cost of developing a pre-pandemic H2N2 vaccine. They cite the impediments to distributing vaccines internationally (including regulatory hurdles); the degree of public distrust of vaccines, particularly in Europe and the United States; and the finite public-health resources available. However, another major influenza pandemic is likely to cost far more and create a much greater health burden than a well-planned pre-emptive programme. The US Centers for Disease Control and Prevention estimates that a pandemic outbreak costs the United States between \$71 billion and \$167 billion.

Sceptics also raise concerns over whether it makes sense to expose individuals to vaccines for pathogens that are not currently in circulation. Fortunately, the previously licensed H2N2 vaccine has

VULNERABILITY OF YOUTH

Vaccination against H2N2 ended in the late 1960s, so people younger than 50 have little immunity to the virus (mean antibody counts shown with 95% confidence intervals).



a proven safety and efficacy record. The virus that does emerge might have evolved to the point of being able to evade human immunity to this vaccine, but the H2 surface protein's high degree of conservation suggests that this is unlikely.

With the knowledge and technologies available today, the efficacy of an H2N2 vaccine could markedly improve, and its cost may decrease in the coming years. A first step towards a pre-emptive vaccine would be to re-examine the safety and efficacy of the existing H2N2 vaccine in animal models and in phase I and II clinical trials. These studies would allow researchers and public-health officials to determine what dose is needed, and establish who to immunize and when.

Our understanding of influenza-virus biology should allow us to prepare for and even mitigate future pandemics. A deliberate and thoughtful strategy for H2N2 vaccination will save lives and spare the world a major public-health crisis. ■

Gary J. Nabel, Chih-Jen Wei and Julie E. Ledgerwood are at the Vaccine Research Center of the National Institutes of Health in Bethesda, Maryland 20892, USA.
e-mail: glabel@nih.gov

The views expressed are those of the authors and do not represent the policy of the National Institutes of Health or the US Department of Health and Human Services.

1. Krause, J. C. *et al.* *J. Virol.* **84**, 3127–3130 (2010).
2. Manicassamy, B. *et al.* *PLoS Pathog.* **6**, e1000745 (2010).
3. Wei, C. J. *et al.* *Sci. Transl. Med.* **2**, 24ra21 (2010).
4. Xu, R. *et al.* *Science* **328**, 357–360 (2010).
5. Chen, G. L. *et al.* *J. Virol.* **84**, 7695–7702 (2010).
6. Hilleman, M. R. *Vaccine* **20**, 3068–3087 (2002).
7. Webster, R. G. *J. Infect. Dis.* **176**, S14–S19 (1997).
8. Ma, W. *et al.* *Proc. Natl Acad. Sci. USA* **104**, 20949–20954 (2007).
9. Kaverin, N. V. *et al.* *Arch. Virol.* **145**, 1059–1066 (2000).
10. Stohr, K. *Nature* **465**, 161 (2010).



From 1957 to 1968, H2N2 influenza killed more than 1 million people.



Biobanks need publicity

Most Europeans haven't heard of their nation's repositories of human blood and tissue samples. Promote them, say **George Gaskell** and **Herbert Gottweis**, or they could fail.

In the late 1990s, the company deCODE Genetics in Reykjavik sought to pool medical and genealogical records with genetic data for the entire Icelandic population — with the intention of marketing the database to scientists. The project quickly came under fire as critics found holes in the company's plans for encrypting data, and in the rules that would govern access, patient consent and data ownership. The firm eventually moved ahead, but filed for bankruptcy in 2009.

All sorts of stumbling blocks — not least the global financial crisis — contributed to the company's demise. A key factor was that by 2003, some 20,000 Icelanders, many of whom had concerns about privacy, had opted out of the database¹.

Biobanks — collections of blood or

tissue samples, often combined with medical, genealogical and lifestyle information — are a powerful tool in the study of complex diseases. In general, tens of thousands of samples are needed to detect associations between genetic factors and a particular disease, so their effectiveness depends on enrolling enough people². Also, biobanks need donors who are prepared to give broad consent to have their samples used in a wide variety of research, because it would be hugely costly and technically onerous to draw up 'narrow' consent forms for each research project.

It is alarming then that two-thirds of Europeans recently surveyed have never heard of biobanks. This was a finding of our research based on the 2010 Eurobarometer on biotechnology³, which canvassed 15,600 people from 32 countries on the issue, and on

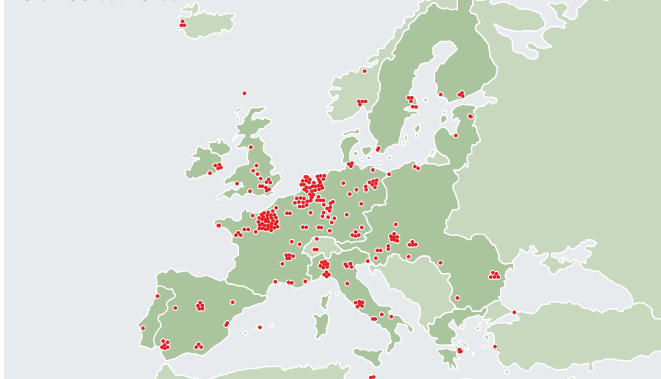
focus groups conducted in seven countries⁴.

Our data provide strong evidence that people are more likely to embrace the idea of biobank research if they are aware of the existence and purposes of the resources. We find that fewer than 2% have actively sought information about biobanks or have discussed them with others frequently. Crucially, only 10% of those who have never heard of biobanks say they would definitely participate, compared with 28% of those who have taken an active interest. And the 'actively interested' are more likely to endorse broad consent — 33% compared with 20% of those unaware of biobanks.

Europe has been at the forefront of the biobank movement. The Biobanking and Biomolecular Resources Research Infrastructure (BBMRI), for instance, launched

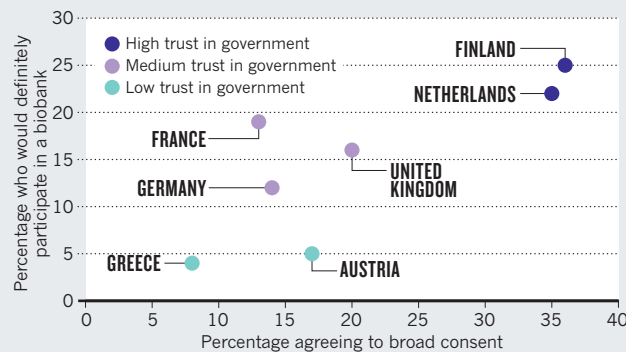
EUROPE'S BIOBANKS

The Biobanking and Biomolecular Resources Research Infrastructure includes in excess of 280 organizations (mainly biobanks) in more than 30 countries.



NATIONAL ATTITUDES

The number of people who would be prepared to participate in biobank research varies greatly by country, owing in part to the trust the citizens place in their government.



in 2007 to enhance cooperation, now involves 270 organizations from 33 countries. The expansion of biobanks in Europe (see 'Europe's biobanks') and elsewhere will deliver little, however, unless potential donors know enough to trust the organizations and the researchers who use them.

TRUST AND CONSENT

The results of our survey suggest that attitudes vary widely across Europe. Some of these differences offer clues as to how public acceptance could be improved.

In northern Europe, people are relatively well informed (80% have heard of biobanks in Iceland for instance, 75% in Sweden and 65% in Norway)³. People in these countries are also more willing to be donors and to give broad consent (see 'National attitudes'). These positive attitudes correspond to strong trust in government, good welfare systems and long traditions of biological samples in these countries. Finland, for example, began research dedicated to establishing national health risk factors in the 1950s⁴. DNA samples have been collected since the 1980s in various projects — such as the Helsinki Sudden Death Study, launched to study the lifestyle and genetic factors predisposing certain Finnish men to this condition.

France, Germany and the United Kingdom form another cluster. Here, people have less trust in the government, moderate awareness of biobanks, show moderate willingness to participate and prefer narrow consent agreements. One participant in a UK focus group summed up a common concern when he asked, "Do I trust what you're going to do with these samples?". People from these countries also think that biobanks should offer benefits to donors: information about cholesterol levels and other health risks, for example.

Austria and Greece form a third cluster, characterized by the lowest levels of willingness to participate: only 5% of Austrians and 4% of Greeks say they would 'definitely'

contribute samples. These people are the most concerned about privacy and about what their genetic data might be used for. Greece also had the highest number of people selecting 'narrow consent'.

In our focus groups, Greek participants explained that they don't trust their country's political system and therefore worry that the data would fall into the wrong hands — the police, insurance companies and the government. In Austria, caution about biobanks might reflect a lack of enthusiasm for technologies in general and for genetic technologies in particular³.

So how might public perception be improved? Counter-intuitively, one way to generate interest in and knowledge about biobanks seems to be controversy. Today, the best informed and most enthusiastic public in Europe is in Iceland, which surely has something to do with the deCODE Genetics furore. Similarly, Estonians are relatively positive about biobank research. This is despite heated debates over the public-private partnership proposed to underpin the Estonian Genome Project, leading to the biobank's near collapse in 2003.

SHAPING ATTITUDES

So, biobank organizers should not resist engaging in public discussion for fear that it may be critical. Controversies don't seem to lead people to reject the idea of biobank research per se. Instead they facilitate the spread of information, and improve understanding and sharing of views on what is appropriate and acceptable use of samples.

Researchers have little influence over patterns of trust in government, but they can shape attitudes to a particular biotechnology. They can, for instance: adopt clear regulations and guidelines for research; make goals and funding priorities transparent; and listen to the public and respond to its concerns. Such issues are common to much genomics research⁶.

Online communication tools such as

Facebook should be harnessed to reach doctors, teachers, journalists and minority-group representatives. Efforts could be made to encourage existing donors to spread the message through their social networks.

Science communication must go beyond the simple dissemination of basic information. What is needed is a dialogue with the public, to explain the purposes of biobanks and how they operate, and to give people an opportunity to voice their concerns and conditions for their support and participation. The involvement of the public in forums representing patients' perspectives in the BBMRI project, or in online public consultations and public meetings as planned for the pre-recruitment phase of a Tasmanian biobank in Australia, offer promising models. Such activities need funding, but the cost is trivial compared with the many millions of dollars needed to establish and maintain a biobank.

We will only see the benefits of the estimated US\$1 billion already spent on establishing biobanks worldwide if the general public is actively brought into the picture. ■

George Gaskell is at the London School of Economics and Political Science, London WC2A 2AE, UK. He coordinates a European Commission project on technologies and public ethics (www.stepe.eu). **Herbert Gottweis** is at the Life Science Governance Research Platform, Department of Political Science, University of Vienna, Vienna, Austria.
e-mails: g.gaskell@lse.ac.uk; herbert.gottweis@univie.ac.at

1. Pálsson, G. in *Biobanks: Governance in Comparative Perspective* (eds Gottweis, H. & Petersen, A.) 41–55 (Routledge, 2008).
2. Smith, G. D. et al. *Lancet* **366**, 1484–1498 (2005).
3. Gaskell, G. et al. *Europeans and Biotechnology in 2010: Winds of Change?* (European Commission, 2010); available at <http://go.nature.com/8k3ehj>
4. <http://go.nature.com/6hzchd>
5. Tupasela, A. *Consent Practices and Biomedical Knowledge Production in Tissue Economies* (2008); available at <http://go.nature.com/qvqgdu>
6. Trinidad, S. B. et al. *Science* **331**, 287–288 (2011).



The DeWitts worked together on the unresolved problem of quantum gravity for more than 50 years.

BIOGRAPHY

Canonical prophet

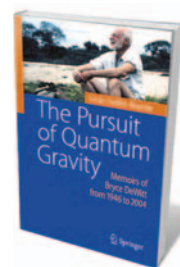
Pedro Ferreira is inspired by a treasury of recollections of physicist Bryce DeWitt by his widow Cécile.

The search for a theory of quantum gravity — the marriage of quantum mechanics and general relativity — is the ultimate frontier of physics. It remains unfinished business, and its creative ferment has attracted many vibrant people, not least Bryce DeWitt, one of the founders of the field. In *The Pursuit of Quantum Gravity*, his widow and long-time research collaborator Cécile DeWitt-Morette offers a personal view of the great man.

This is not a typical memoir, but a fascinating collage of essays, letters and commentary that is like a treasure chest: a stack of priceless, dusty heirlooms.

Although the digressions into obscure aspects of quantum gravity will challenge some readers, the erratic narrative gradually builds up an impressionistic tableau of DeWitt's remarkable journey through postwar physics.

DeWitt's interest in quantum gravity began in the late 1940s when he was studying quantum field theory with Julian Schwinger at Harvard University in Cambridge, Massachusetts. Struck by gravity's "splendid isolation" in the canon of fundamental forces, DeWitt asked: "What if one simply dragged it forcibly into the mainstream of theoretical physics and quantized it?" His search



The Pursuit of Quantum Gravity: Memoirs of Bryce DeWitt from 1946 to 2004

CÉCILE DEWITT-MORETTE
Springer: 2011.
147 pp. \$49.95

for a means of doing so became his lifelong quest.

A trio of papers that DeWitt published in 1967 is at the heart of the field. The Trilogy, as it is known, is lucid, visionary and packed with results. The papers cover two approaches to quantum gravity. The first publication describes the 'canonical approach' that served up the Wheeler-DeWitt equation (also named after relativist John Wheeler) for the quantum evolution of the geometry of space. The second and third papers develop a 'covariant approach' to quantizing gravity, which considers the sum of all the possible configurations a system can have in space and time, and is an extension of a method that Richard Feynman applied to quantum mechanics and quantum field theory.

DeWitt, the book reveals, much preferred the covariant approach, openly ruining the day he came up with the Wheeler-DeWitt equation, which treated space and time separately rather than as a combination. His work on the covariant approach went on to play a key part in quantizing the electromagnetic, weak and strong forces during the 1960s and 1970s. The canonical approach also influenced Stephen Hawking and others in their proposal of a "wave function of the Universe".

The Trilogy was the sacred creed of quantum gravity for decades, and DeWitt was the prophet. He was a biblical character in many ways. Raised in California in a devoutly religious family (although he kept his distance from religion in later life), DeWitt had the towering, stern presence of a church elder. When he walked into a room, backs straightened. He would leave a seminar if he thought it was not worth his while. He had no time for sloppiness; things had to be done properly. Once ideas were published, they were set in stone.

DeWitt loved to travel, considering himself a "space traveller". He enlisted as a US Navy pilot in January 1944 but did not see any combat. Following his graduate studies in Harvard during the late 1940s, he hopped around the globe, working at Princeton University, at the Swiss Federal Institute of Technology in Zurich and, in what Steven Weinberg judged a "sojourn

NATURE.COM

For a biography of physicist Hugh Everett III, see:
go.nature.com/4zk9oz

[that] did not make good professional sense, but ... suited his roving spirit", at the Tata Institute of Fundamental Research in Mumbai, India. In 1973 he led an expedition to Mauritania to repeat Arthur Eddington's 1919 measurement of the gravitational bending of starlight around the Sun's limb during a solar eclipse. Cécile and Bryce were also known for taking long walks in the stunning landscape of the American West and in the Alps, where Cécile set up the Les Houches scientific lecture series.

DeWitt's academic career was colourful and varied. After his postdoctoral fellowships, he worked in California from 1952 to 1955 at the Lawrence Livermore Laboratory with Edward Teller on modelling nuclear artillery shells. He briefly led the Institute of Field Physics at Chapel Hill, North Carolina, a centre funded by the industrialist Agnew Bahnson, who hoped to unlock technology from basic research in gravity. In 1973 DeWitt ended up, along with Cécile, at the Center for Relativity at the University of Texas at Austin, which was at the heart of the revival of research interest in general relativity in the 1960s.

Sadly lacking from the book is a personal description of the DeWitts' epic love story. Bryce and Cécile were together for more than 50 years and had four children. According to the publication list, the couple wrote their first paper together in 1952 and their last in 2004. Yet we learn little about their joint lives. There are scraps of information: Bryce mentions that he "cast glances" at Cécile in 1949, when he met her at the Institute for Advanced Studies in Princeton. A picture of Bryce and Cécile is captioned with a quote from one of DeWitt's obituaries: "They became lifelong sparring partners." I would have liked to have known more.

Nevertheless, the book is enthralling. It made me want to take up the challenge of quantum gravity, to follow the extraordinary trail set down by Bryce and Cécile to that untamed frontier of physics. ■

Pedro Ferreira is a professor of astrophysics at the University of Oxford, Keble Road, Oxford OX1 3RH, UK, and author of *The State of the Universe*. e-mail: p.ferreira1@physics.ox.ac.uk

GEOGRAPHY

Islamic views of Earth

A magisterial series revolutionizes our understanding of Arabic geography, finds Celâl Şengör.

Scientific geography was invented by the Greeks in the sixth century BC and reached its apogee in the mid-nineteenth century in Europe. In this five-volume work, *Geschichte des Arabischen Schrifttums* (*History of Arabic Literature*) — the last two of which were published in 2010 — science historian Fuat Sezgin shows that Muslim scholars played an integral part in developing Greek ideas about the planet, and explains how influential their work was in the West.

By reviewing and analysing a vast corpus of Arabic geographical writings from the eighth to the seventeenth centuries, Sezgin revolutionizes our view of the history of geography. The first three volumes, plus an atlas, cover the development of mathematical geography and cartography. The latter two volumes span a range of topics, from the history of general and regional geography, through topography and cosmology to the history of Arabic-Islamic travel reports. He brackets these as 'anthropogeography' to underline their relevance to human existence.

In the first three books, Sezgin establishes that Muslim scholars — starting with geographers during the reign of al-Ma'mun (786–833 AD), a caliph of Abbasid — inherited and developed the geographical and cartographical tradition of the Persian Sassanids and Greeks, from Eratosthenes to Ptolemy. The Muslims improved on Ptolemy's database of longitudes and latitudes of features such as cities, mountain ranges, rivers and continental coastlines.

They initiated geodetic work in the ninth century AD by re-measuring the length of a degree of meridian in the plains of Mosul and Damascus and at Mount Casius, or Jebel Aqra, near Hatay (ancient Antioch) in present-day Turkey, and by fixing the positions of geographical features on the basis of astronomical observations. In the eleventh century, this culminated in Turkish scholar Abu Rayhan al-Biruni's (973–1048) great

Geschichte des Arabischen Schrifttums, Band XIV. Anthropogeographie Teil I: Gesamt- und Ländergeographie; Stadt- und Regionalgeographie

FUAT SEZGIN

J. W. Goethe Univ.: 2010. 553 pp. €124. (In German.)

Geschichte des Arabischen Schrifttums, Band XV. Anthropogeographie Teil II: Topographie, Geographische Lexika, Kosmographie, Kosmologie, Reiseberichte

FUAT SEZGIN

J. W. Goethe Univ.: 2010. 470 pp. €117. (In German.)

work *Measurement of Distances Between Known Places on Earth*.

The many maps produced as a result of these activities by Muslim scholars eventually reached Europe. Beginning with portolan maps in the fourteenth century — predecessors of accurate navigational charts showing coastlines and harbours — they influenced the European cartography of Asia, Africa and the Indian Ocean until well into the nineteenth century.

Sezgin's two most recent volumes go beyond cartography. Arabic geographical activity started before the seventh century AD, he explains. Much geographical information was included in poetry — the names of places, oases and other watering sites. So the first geographical publications in Islamic culture were

penned by philologists in the second century of Islam, following the expansion of the Arabic empire under the Umayyads and then the Abbasids. The first map by a Muslim author — a remarkable map of the world in the form of a bird — was made in the seventh century by a young friend of the prophet Mohammad, 'Abdallah b'Amr b. al-Āṣ.

The demand for broader Muslim geographies grew as European interests expanded beyond its borders from the fourteenth to the sixteenth centuries. Travellers hunted for works on



A world map by ninth-century Muslim geographers.

➔NATURE.COM

For more on Islamic history of science: go.nature.com/gbipbz

F. SEZGIN

[that] did not make good professional sense, but ... suited his roving spirit", at the Tata Institute of Fundamental Research in Mumbai, India. In 1973 he led an expedition to Mauritania to repeat Arthur Eddington's 1919 measurement of the gravitational bending of starlight around the Sun's limb during a solar eclipse. Cécile and Bryce were also known for taking long walks in the stunning landscape of the American West and in the Alps, where Cécile set up the Les Houches scientific lecture series.

DeWitt's academic career was colourful and varied. After his postdoctoral fellowships, he worked in California from 1952 to 1955 at the Lawrence Livermore Laboratory with Edward Teller on modelling nuclear artillery shells. He briefly led the Institute of Field Physics at Chapel Hill, North Carolina, a centre funded by the industrialist Agnew Bahnson, who hoped to unlock technology from basic research in gravity. In 1973 DeWitt ended up, along with Cécile, at the Center for Relativity at the University of Texas at Austin, which was at the heart of the revival of research interest in general relativity in the 1960s.

Sadly lacking from the book is a personal description of the DeWitts' epic love story. Bryce and Cécile were together for more than 50 years and had four children. According to the publication list, the couple wrote their first paper together in 1952 and their last in 2004. Yet we learn little about their joint lives. There are scraps of information: Bryce mentions that he "cast glances" at Cécile in 1949, when he met her at the Institute for Advanced Studies in Princeton. A picture of Bryce and Cécile is captioned with a quote from one of DeWitt's obituaries: "They became lifelong sparring partners." I would have liked to have known more.

Nevertheless, the book is enthralling. It made me want to take up the challenge of quantum gravity, to follow the extraordinary trail set down by Bryce and Cécile to that untamed frontier of physics. ■

Pedro Ferreira is a professor of astrophysics at the University of Oxford, Keble Road, Oxford OX1 3RH, UK, and author of *The State of the Universe*. e-mail: p.ferreira1@physics.ox.ac.uk

GEOGRAPHY

Islamic views of Earth

A magisterial series revolutionizes our understanding of Arabic geography, finds Celâl Şengör.

Scientific geography was invented by the Greeks in the sixth century BC and reached its apogee in the mid-nineteenth century in Europe. In this five-volume work, *Geschichte des Arabischen Schrifttums* (*History of Arabic Literature*) — the last two of which were published in 2010 — science historian Fuat Sezgin shows that Muslim scholars played an integral part in developing Greek ideas about the planet, and explains how influential their work was in the West.

By reviewing and analysing a vast corpus of Arabic geographical writings from the eighth to the seventeenth centuries, Sezgin revolutionizes our view of the history of geography. The first three volumes, plus an atlas, cover the development of mathematical geography and cartography. The latter two volumes span a range of topics, from the history of general and regional geography, through topography and cosmology to the history of Arabic-Islamic travel reports. He brackets these as 'anthropogeography' to underline their relevance to human existence.

In the first three books, Sezgin establishes that Muslim scholars — starting with geographers during the reign of al-Ma'mun (786–833 AD), a caliph of Abbasid — inherited and developed the geographical and cartographical tradition of the Persian Sassanids and Greeks, from Eratosthenes to Ptolemy. The Muslims improved on Ptolemy's database of longitudes and latitudes of features such as cities, mountain ranges, rivers and continental coastlines.

They initiated geodetic work in the ninth century AD by re-measuring the length of a degree of meridian in the plains of Mosul and Damascus and at Mount Casius, or Jebel Aqra, near Hatay (ancient Antioch) in present-day Turkey, and by fixing the positions of geographical features on the basis of astronomical observations. In the eleventh century, this culminated in Turkish scholar Abu Rayhan al-Biruni's (973–1048) great

Geschichte des Arabischen Schrifttums, Band XIV. Anthropogeographie Teil I: Gesamt- und Ländergeographie; Stadt- und Regionalgeographie

FUAT SEZGIN

J. W. Goethe Univ.: 2010. 553 pp. €124. (In German.)

Geschichte des Arabischen Schrifttums, Band XV. Anthropogeographie Teil II: Topographie, Geographische Lexika, Kosmographie, Kosmologie, Reiseberichte

FUAT SEZGIN

J. W. Goethe Univ.: 2010. 470 pp. €117. (In German.)

work *Measurement of Distances Between Known Places on Earth*.

The many maps produced as a result of these activities by Muslim scholars eventually reached Europe. Beginning with portolan maps in the fourteenth century — predecessors of accurate navigational charts showing coastlines and harbours — they influenced the European cartography of Asia, Africa and the Indian Ocean until well into the nineteenth century.

Sezgin's two most recent volumes go beyond cartography. Arabic geographical activity started before the seventh century AD, he explains. Much geographical information was included in poetry — the names of places, oases and other watering sites. So the first geographical publications in Islamic culture were

penned by philologists in the second century of Islam, following the expansion of the Arabic empire under the Umayyads and then the Abbasids. The first map by a Muslim author — a remarkable map of the world in the form of a bird — was made in the seventh century by a young friend of the prophet Mohammad, 'Abdallah b'Amr b. al-Āṣ.

The demand for broader Muslim geographies grew as European interests expanded beyond its borders from the fourteenth to the sixteenth centuries. Travellers hunted for works on



A world map by ninth-century Muslim geographers.

➔NATURE.COM

For more on Islamic history of science: go.nature.com/gbipbz

F. SEZGIN

geography as well as those on the economics, vegetation, fauna and statistics of regions, countries and cities. Interest in archaeology and the study of ancient inscriptions also increased in Europe because those subjects were widely covered by Muslim authors.

The first Muslim book of general geography to be published in Arabic in Europe (in Rome in 1585) was *Garden of Strange Things of the Earth and Lands* by a previously unknown Ottoman author called Sālāmī b. Gündoğdi aş-Şālihi. Its influence on European knowledge cannot have been great, because it was never translated. The second such book, by contrast, had an immense influence. Translated into Latin in abridged form in 1619 under the misleading title of *Geographia Nubiensis*, the book was al-Idrisi's famous *Journey of Those who are Amazed*, written for the Norman king Roger II in Sicily in the twelfth century. It accompanied

al-Idrisi's world map, the Tabula Rogeriana, showing a dagger-shaped Africa and the confluence of the Indian and the Atlantic Oceans.

Sezgin ends his review in the seventeenth century, when European geographical activities began to overtake the Muslim effort. He singles out for praise the huge efforts of the European orientalists. From the late eighteenth century onwards, they unearthed huge amounts of material from the libraries and book shops of Islamic countries, transporting it to European libraries for study. This was despite the devastation of the Crusades that began in the eleventh century, the Mongol invasions of the thirteenth century and subsequent years of neglect. Sezgin's eulogy stands in contrast to the attack on the orientalists by Palestinian-American cultural critic Edward Said, who called them sinister servants of imperial powers.

The importance of Sezgin's series cannot

be overestimated. It opens up a little-known world, much of which was forgotten even in Muslim countries. And it contains some unexpected gems, including the best assessment of German geographer Carl Ritter's 20-volume work on comparative geography (1822–59), often misrepresented as a book of history with a teleological bent. Sezgin's books also confirm the intellectual place of seventeenth-century Turkish travel writer Evliya Çelebi as an original thinker and observer in the Ottoman world, a fact that is under-appreciated.

The first two volumes of Sezgin's series are translated into English; I hope that the rest will soon follow. Sezgin's immense scholarship deserves a wide readership. ■

A. M. Celâl Şengör is professor of geology at the Istanbul Technical University, Ayazaga 34469 Istanbul, Turkey.
e-mail: sengor@itu.edu.tr

FICTION

Attack of the killer fungi

Philip Ball applauds physicist Paul McEuen's debut thriller about a madness-inducing mould.

One of my more humdrum obligations as a science commentator was to read Michael Crichton's *Prey*, his 2002 thriller based on the premise of nanotechnological robot swarms run amok. As a novice in this genre, I found myself comparing his characters' psychological implausibility to the illogical quirks of figures from myth and legend. But with guns.

Crichton made millions with his formula; *Spiral* deserves to do the same for Paul McEuen, a physicist at Cornell University in New York. His debut novel is more enjoyable and more palatable than Crichton's and boasts impeccable science.

Even so, nothing in *Spiral* bucks the thriller formula. Every scene is tailored for the screen, and the film rights have already been sold. The dialogue reflects how people speak in blockbusters, not in real life, and the story has the familiar cast: the vulnerable but plucky mother, the ruthless assassin, the sadistic billionaire, the child in peril, and so on. There's the race against time, the apocalyptic threat. And, just as films like this offer a great ride when done well, so too does *Spiral*.

The fictional tale begins at the end of the Second World War, when young Irish microbiologist Liam Connor is brought on board a US warship to witness the effects of a devastating biological weapon developed by the Japanese: a fungal infection called the

Uzumaki that induces hallucinations and madness, and is ultimately fatal. Connor ends up hiding away a tiny vial of the stuff, wrestled from the Japanese engineer Hitoshi Kitano who was responsible for developing it.

Sixty years later, Connor is an octogenarian with a Nobel prize, and still in active research at Cornell. Unknown to the authorities, he has for decades been secretly searching for the cure that he is sure will one day be needed for the Uzumaki. Aware that a cure would turn the deadly fungus into a potential weapon by conferring protection only on some, he is determined to keep his work from the US military. Then he is found dead at the bottom of a gorge, apparently having thrown himself off a bridge to escape from a mysterious woman caught on security cameras. His coded last message to his colleague Jake Sterling, his granddaughter Maggie and her son Dylan makes them the only people who can prevent a global outbreak of the killer fungus. But who is behind the fiendish scheme to release it?

You can see a lot of the plot coming — the denouement even involves the old chestnut of who gets to the gun first. But that doesn't detract from the page-turning quality. It is a delight to see how McEuen — an expert in carbon-nanotube physics and nano-electronics — has marshalled his knowledge to kit out the technical plot devices. Nanotech-



Spiral: A Novel
PAUL MCEUEN
Dial Press/Headline:
2011. 320 pp.
\$25/£19.99

nology, microbiology, information technology and synthetic biology are all brought into play in a convincing, unforced manner. Devotees of scientific trends will recognize many elements, from genetically engineered oscillating fluorescence to microfluidic labs-on-chips.

McEuen shows that the imagination of an inventive scientist is far more interesting than that of a writer who has merely done his homework. In his use of science he trumps Crichton and many other novelists who like to spice their narratives with cutting-edge science. I confess that my interest finds less purchase with square-jawed, stolid heroes like Jake, whose physical prowess and ex-army credentials are carefully established in preparation for gutsy displays. But that is the genre, and Jake is less tiresomely bland than the wooden leads in the books of Dan Brown and Crichton.

A more appealing hero is Cornell University itself, which enjoys a touching love letter here from the author. But, as ever, the stars of the show are the villains: the microcrawlers that scabble ominously across the book's cover — microelectromechanical devices with a seriously bad attitude.

Next time, McEuen should allow himself to push harder at the genre's boundaries. And I do hope there will be a next time, if he can escape the lab bench and the jaws of Hollywood. ■

Philip Ball is a writer based in London. His latest book is *Unnatural*.

geography as well as those on the economics, vegetation, fauna and statistics of regions, countries and cities. Interest in archaeology and the study of ancient inscriptions also increased in Europe because those subjects were widely covered by Muslim authors.

The first Muslim book of general geography to be published in Arabic in Europe (in Rome in 1585) was *Garden of Strange Things of the Earth and Lands* by a previously unknown Ottoman author called Sālāmī b. Gündoğdi aş-Şālihi. Its influence on European knowledge cannot have been great, because it was never translated. The second such book, by contrast, had an immense influence. Translated into Latin in abridged form in 1619 under the misleading title of *Geographia Nubiensis*, the book was al-Idrisi's famous *Journey of Those who are Amazed*, written for the Norman king Roger II in Sicily in the twelfth century. It accompanied

al-Idrisi's world map, the Tabula Rogeriana, showing a dagger-shaped Africa and the confluence of the Indian and the Atlantic Oceans.

Sezgin ends his review in the seventeenth century, when European geographical activities began to overtake the Muslim effort. He singles out for praise the huge efforts of the European orientalists. From the late eighteenth century onwards, they unearthed huge amounts of material from the libraries and book shops of Islamic countries, transporting it to European libraries for study. This was despite the devastation of the Crusades that began in the eleventh century, the Mongol invasions of the thirteenth century and subsequent years of neglect. Sezgin's eulogy stands in contrast to the attack on the orientalists by Palestinian-American cultural critic Edward Said, who called them sinister servants of imperial powers.

The importance of Sezgin's series cannot

be overestimated. It opens up a little-known world, much of which was forgotten even in Muslim countries. And it contains some unexpected gems, including the best assessment of German geographer Carl Ritter's 20-volume work on comparative geography (1822–59), often misrepresented as a book of history with a teleological bent. Sezgin's books also confirm the intellectual place of seventeenth-century Turkish travel writer Evliya Çelebi as an original thinker and observer in the Ottoman world, a fact that is under-appreciated.

The first two volumes of Sezgin's series are translated into English; I hope that the rest will soon follow. Sezgin's immense scholarship deserves a wide readership. ■

A. M. Celâl Şengör is professor of geology at the Istanbul Technical University, Ayazaga 34469 Istanbul, Turkey.
e-mail: sengor@itu.edu.tr

FICTION

Attack of the killer fungi

Philip Ball applauds physicist Paul McEuen's debut thriller about a madness-inducing mould.

One of my more humdrum obligations as a science commentator was to read Michael Crichton's *Prey*, his 2002 thriller based on the premise of nanotechnological robot swarms run amok. As a novice in this genre, I found myself comparing his characters' psychological implausibility to the illogical quirks of figures from myth and legend. But with guns.

Crichton made millions with his formula; *Spiral* deserves to do the same for Paul McEuen, a physicist at Cornell University in New York. His debut novel is more enjoyable and more palatable than Crichton's and boasts impeccable science.

Even so, nothing in *Spiral* bucks the thriller formula. Every scene is tailored for the screen, and the film rights have already been sold. The dialogue reflects how people speak in blockbusters, not in real life, and the story has the familiar cast: the vulnerable but plucky mother, the ruthless assassin, the sadistic billionaire, the child in peril, and so on. There's the race against time, the apocalyptic threat. And, just as films like this offer a great ride when done well, so too does *Spiral*.

The fictional tale begins at the end of the Second World War, when young Irish microbiologist Liam Connor is brought on board a US warship to witness the effects of a devastating biological weapon developed by the Japanese: a fungal infection called the

Uzumaki that induces hallucinations and madness, and is ultimately fatal. Connor ends up hiding away a tiny vial of the stuff, wrestled from the Japanese engineer Hitoshi Kitano who was responsible for developing it.

Sixty years later, Connor is an octogenarian with a Nobel prize, and still in active research at Cornell. Unknown to the authorities, he has for decades been secretly searching for the cure that he is sure will one day be needed for the Uzumaki. Aware that a cure would turn the deadly fungus into a potential weapon by conferring protection only on some, he is determined to keep his work from the US military. Then he is found dead at the bottom of a gorge, apparently having thrown himself off a bridge to escape from a mysterious woman caught on security cameras. His coded last message to his colleague Jake Sterling, his granddaughter Maggie and her son Dylan makes them the only people who can prevent a global outbreak of the killer fungus. But who is behind the fiendish scheme to release it?

You can see a lot of the plot coming — the denouement even involves the old chestnut of who gets to the gun first. But that doesn't detract from the page-turning quality. It is a delight to see how McEuen — an expert in carbon-nanotube physics and nano-electronics — has marshalled his knowledge to kit out the technical plot devices. Nanotech-



Spiral: A Novel
PAUL MCEUEN
Dial Press/Headline:
2011. 320 pp.
\$25/£19.99

nology, microbiology, information technology and synthetic biology are all brought into play in a convincing, unforced manner. Devotees of scientific trends will recognize many elements, from genetically engineered oscillating fluorescence to microfluidic labs-on-chips.

McEuen shows that the imagination of an inventive scientist is far more interesting than that of a writer who has merely done his homework. In his use of science he trumps Crichton and many other novelists who like to spice their narratives with cutting-edge science. I confess that my interest finds less purchase with square-jawed, stolid heroes like Jake, whose physical prowess and ex-army credentials are carefully established in preparation for gutsy displays. But that is the genre, and Jake is less tiresomely bland than the wooden leads in the books of Dan Brown and Crichton.

A more appealing hero is Cornell University itself, which enjoys a touching love letter here from the author. But, as ever, the stars of the show are the villains: the microcrawlers that scabble ominously across the book's cover — microelectromechanical devices with a seriously bad attitude.

Next time, McEuen should allow himself to push harder at the genre's boundaries. And I do hope there will be a next time, if he can escape the lab bench and the jaws of Hollywood. ■

Philip Ball is a writer based in London. His latest book is *Unnatural*.



University College London's Grant Museum reopens this week in a new home with more space for specimens.

MUSEUMS

Campus treasures

Sally MacDonald and Jack Ashby explain how best to air the gems hidden within university science collections.

University museums contain some of the richest and most extraordinary collections in the world, from Charles Darwin's Galapagos finches at the University of Cambridge, UK, to Louis Agassiz's vertebrate palaeontology specimens at Harvard University in Massachusetts. Such collections can draw on excellent resources — academic minds, vast libraries and student enthusiasm. But now that universities are increasingly required to demonstrate their impact on society, their museums must also claim a new role as vital public spaces.

Aiming to do just that, the Grant Museum of Zoology at University College London (UCL) reopens to the public this week after an eight-month closure while it moved to a new building. With the extra space in its new home, as its directors we intend to make it a vibrant place for experiment and dialogue by offering provocative, interactive and regularly changing displays. The museum retains its Victorian cabinets, crammed with skeletons and specimens in jars, with many coming out of storage for the first time — including dozens of primate and carnivore skeletons, newly rediscovered dodo material and a rare pterosaur fossil. But we will also pilot new ways of engaging visitors with science and scientists.

There is a pressing need to review the purpose of university museums. In the nineteenth and early twentieth centuries, when many significant collections were founded, a museum was seen as an essential teaching

Grant Museum of Zoology
Rockefeller Building,
University College
London, UK.
Opens 15 March 2011.

Robert Edmond Grant, who held the first chair in zoology and comparative anatomy at UCL. It has one of only seven known quagga skeletons (the extinct South African zebra, considered the rarest skeleton in the world) and a singular adult fluid-preserved specimen of a thylacine, or Tasmanian tiger, which was dissected by Thomas Huxley.

The Grant Museum has been well maintained, but many other university collections are in a parlous state — uncurated and unused, consigned to cramped basements or, on occasion, to the bin. Poor advocacy of their role within their parent institution (or beyond) is one reason for the lack of funding. Pressure on space is another: campus museums must compete with income-generating research facilities and student amenities. And whereas disciplines such as archaeology and art history use collections extensively, others, including the life sciences, have turned to digital resources instead of the real things.

Natural history collections have been particularly affected by this change in academic fashions, with the result that most science-engagement funding has flowed into science centres — concept-driven and

resource for disciplines ranging from archaeology to zoology. The Grant Museum, for example, was set up in 1827 to house the growing collection of

specimen-free. With the financial situation now bleaker than ever, university museums must again reconnect with research and teaching and open their doors to the public.

Some university museums have already become major science-engagement venues for their cities. The Grant Museum first opened to the public in 1996, and offers a lively natural history learning programme, including lectures, film nights, debates and treasure hunts. The Peabody Museum of Natural History at Yale University in New Haven, Connecticut, which holds around 11 million objects ranging from meteorites and insects to scientific instruments, runs many public workshops. The 30 million objects, including thousands of type specimens, at the Humboldt University Museum of Natural History in Berlin are used in hands-on workshops and changing exhibitions.

RESEARCH RELEVANT

To remain relevant to teaching and research, the potential of historic collections needs continual reassessment. Although teaching in the life sciences has shifted away from whole organisms to a more molecular focus, natural history collections provide historical snapshots of past biodiversity. Biologists can analyse skeletal material, using new technologies such as three-dimensional scanners. Fossils and modern specimens are accessible to DNA analysis for studies in phylogenetics, population dynamics and conservation biology. At the Grant Museum, genetic samples have been used to date the extinction of the giant Irish elk, *Megaloceros giganteus*.

Successful university museums must promote themselves to academics and students from a wide range of disciplines. The Grant Museum provides object-based classes that go far beyond biology. We support teaching in arts management, archaeology, architecture, museum studies and even the performing arts — dance students from the Royal Academy of Dramatic Art use the museum to learn the anatomical basis of animal movement so that they can recreate it. Specimen-based modules run on the university's virtual-learning environment, allowing instantaneous feedback on students' coursework and enhancing their observational skills. The collection database has been available online since 2008.

There are legal and ethical challenges to working with old collections in the public domain. No nineteenth-century assemblage of primates was complete without a human, but the storage and use of human remains is now tightly regulated; in Britain, no human material less than 100 years old can be displayed or used in teaching without an expensive licence. Uncertainties

NATURE.COM
For a Q&A with
museum designer
Edwin Schlossberg:
go.nature.com/zzu2nc

about acquisition dates raise the question of whether museums should risk displaying such remains, or use replicas instead. We have repatriated human remains to communities in Australia and New Zealand and are in the process of transferring all unprovenanced remains to a licensed teaching collection.

The display of specimens of endangered species also raises problems. The Convention on International Trade in Endangered Species restricts the commercial use of such material, and this includes their use in museums or one-off events for which admission is charged, or the featuring of listed animals in filming. We must carefully consider which specimens are used for such events.

We hope that the Grant Museum will become a space where the public, students and academics alike can use objects to tackle big questions in the life sciences and engage with the way museums work. Interactive labels, accessed on iPads located next to the exhibits, pose provocative questions, such as whether animal and human remains should be treated any differently. Visitors can respond using the iPads or smartphones, Twitter or home computers. Questions may be crowd-sourced in future, opening up the potential for the public to curate museum displays.

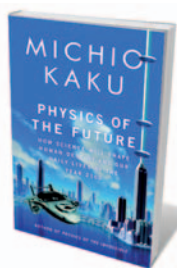
Like our much larger cousins, London's Natural History Museum or the American Museum of Natural History in New York, the Grant Museum is a working institution, where the collections are accessed constantly and real science is done. University research will be showcased by encouraging artists and scientists to co-curate designated cases or art installations using the collections. And discussions, teaching and debates will take place in a space in the centre of the displays. The museum staff will also be accessible so that visitors can ask questions.

The redisplay of the Grant Museum had to be done quickly and cheaply in comparison to typical museum redisplays: UCL footed the £326,000 (US\$530,000) bill. This imposed constraints but also brought benefits: no need to satisfy external funders, no problem with last-minute changes. This museum-cum-lab is a work in progress and we will no doubt struggle with stretched budgets. But if we can continue to find new ways to engage researchers, students and visitors with the collections, the Grant Museum should continue to evolve and thrive, at 184 years old and counting. ■

Sally MacDonald is director of UCL museums, collections and public engagement at University College London, Gower Street, London WC1E 6BT, UK.
e-mail: tcrrnsmd@ucl.ac.uk

Jack Ashby is head of learning and access at the Grant Museum of Zoology, University College London, UK.
e-mail: j.ashby@ucl.ac.uk

Books in brief



Physics of the Future: How Science Will Shape Human Destiny and Our Daily Lives by the Year 2100

Michio Kaku DOUBLEDAY 416 pp. \$28.95 (2011)

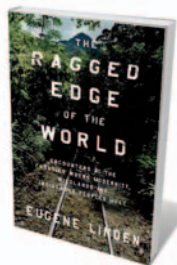
What will the world be like in the year 2100? Full of new life forms, emotional robots and antimatter rockets, according to string theorist and author Michio Kaku. Limiting himself to extensions of current technology, he synthesizes the ideas of 300 experts in diverse fields from neuroscience to nanotechnology. Whereas researchers expect predictions such as limitless computing and synthetic organisms to be fulfilled, others, such as prolific space travel, will be harder to pull off.



Cricket Radio: Tuning In the Night-Singing Insects

John Himmelman BELKNAP PRESS 272 pp. \$22.95 (2011)

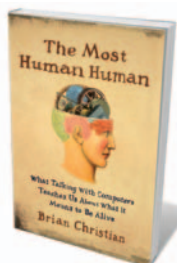
Warm summer nights resonate with the chirruping of crickets and other night-singing insects. Amateur naturalist and writer John Himmelman seeks to reconnect us to these often-unseen musical creatures by mingling tales of his own searches for grasshoppers, cicadas and katydids with the latest professional research investigating why and how they sing, to attract mates and ward off predators. Himmelman collects the sounds of cricket calls on a website that accompanies the book (see go.nature.com/czodcg).



The Ragged Edge of the World: Encounters at the Frontier Where Modernity, Wildlands, and Indigenous Peoples Meet

Eugene Linden VIKING 272 pp. \$26.95 (2011)

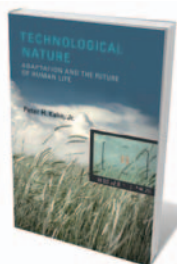
In 40 years of reporting, environmental journalist Eugene Linden has witnessed great changes to swathes of the globe. Dense forests in New Guinea and Borneo have been opened up for timber and minerals; tourism has invaded remote areas such as Machu Picchu in Peru and the Antarctic. Linden calls for the expansion of ecological corridors and 'peace parks' to help species adapt to globalization and combat exploitation of the environment by the increasing human population.



The Most Human Human: What Talking with Computers Teaches Us About What It Means to Be Alive

Brian Christian DOUBLEDAY 320 pp. \$27.95 (2011)

When it comes to good conversation, humans are usually better than a machine. Why artificial intelligence so often fails to live up to our emotional expectations, and what that tells us about being human, is the focus of science writer and poet Brian Christian's book. He relates what it was like to take part in a Turing test, which pitted him against a computer in an anonymous chat with a third-party judge in 2009, and the lengths to which he went to make sure that a human won.



Technological Nature: Adaptation and the Future of Human Life

Peter H. Kahn MIT PRESS 240 pp. \$24.95 (2011)

From watching wildlife documentaries to stalking animals with video cameras, we increasingly interact with nature through technology. After studying people's responses to real-time broadcasts of natural scenes and their play with robotic pets, psychologist Peter Kahn concludes that exposure to digital nature is better than having no natural experience at all. But, he finds, you cannot beat the real thing, arguing that we should view this 'technological nature' as a bonus rather than as a substitute for tangible natural wonders.

about acquisition dates raise the question of whether museums should risk displaying such remains, or use replicas instead. We have repatriated human remains to communities in Australia and New Zealand and are in the process of transferring all unprovenanced remains to a licensed teaching collection.

The display of specimens of endangered species also raises problems. The Convention on International Trade in Endangered Species restricts the commercial use of such material, and this includes their use in museums or one-off events for which admission is charged, or the featuring of listed animals in filming. We must carefully consider which specimens are used for such events.

We hope that the Grant Museum will become a space where the public, students and academics alike can use objects to tackle big questions in the life sciences and engage with the way museums work. Interactive labels, accessed on iPads located next to the exhibits, pose provocative questions, such as whether animal and human remains should be treated any differently. Visitors can respond using the iPads or smartphones, Twitter or home computers. Questions may be crowd-sourced in future, opening up the potential for the public to curate museum displays.

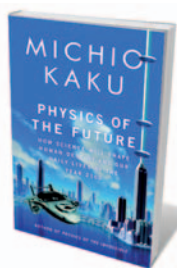
Like our much larger cousins, London's Natural History Museum or the American Museum of Natural History in New York, the Grant Museum is a working institution, where the collections are accessed constantly and real science is done. University research will be showcased by encouraging artists and scientists to co-curate designated cases or art installations using the collections. And discussions, teaching and debates will take place in a space in the centre of the displays. The museum staff will also be accessible so that visitors can ask questions.

The redisplay of the Grant Museum had to be done quickly and cheaply in comparison to typical museum redisplays: UCL footed the £326,000 (US\$530,000) bill. This imposed constraints but also brought benefits: no need to satisfy external funders, no problem with last-minute changes. This museum-cum-lab is a work in progress and we will no doubt struggle with stretched budgets. But if we can continue to find new ways to engage researchers, students and visitors with the collections, the Grant Museum should continue to evolve and thrive, at 184 years old and counting. ■

Sally MacDonald is director of UCL museums, collections and public engagement at University College London, Gower Street, London WC1E 6BT, UK.
e-mail: tcrrnsmd@ucl.ac.uk

Jack Ashby is head of learning and access at the Grant Museum of Zoology, University College London, UK.
e-mail: j.ashby@ucl.ac.uk

Books in brief



Physics of the Future: How Science Will Shape Human Destiny and Our Daily Lives by the Year 2100

Michio Kaku DOUBLEDAY 416 pp. \$28.95 (2011)

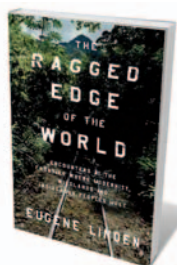
What will the world be like in the year 2100? Full of new life forms, emotional robots and antimatter rockets, according to string theorist and author Michio Kaku. Limiting himself to extensions of current technology, he synthesizes the ideas of 300 experts in diverse fields from neuroscience to nanotechnology. Whereas researchers expect predictions such as limitless computing and synthetic organisms to be fulfilled, others, such as prolific space travel, will be harder to pull off.



Cricket Radio: Tuning In the Night-Singing Insects

John Himmelman BELKNAP PRESS 272 pp. \$22.95 (2011)

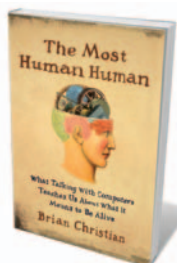
Warm summer nights resonate with the chirruping of crickets and other night-singing insects. Amateur naturalist and writer John Himmelman seeks to reconnect us to these often-unseen musical creatures by mingling tales of his own searches for grasshoppers, cicadas and katydids with the latest professional research investigating why and how they sing, to attract mates and ward off predators. Himmelman collects the sounds of cricket calls on a website that accompanies the book (see go.nature.com/czodcg).



The Ragged Edge of the World: Encounters at the Frontier Where Modernity, Wildlands, and Indigenous Peoples Meet

Eugene Linden VIKING 272 pp. \$26.95 (2011)

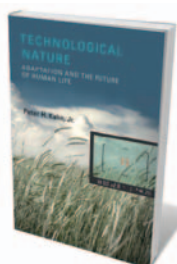
In 40 years of reporting, environmental journalist Eugene Linden has witnessed great changes to swathes of the globe. Dense forests in New Guinea and Borneo have been opened up for timber and minerals; tourism has invaded remote areas such as Machu Picchu in Peru and the Antarctic. Linden calls for the expansion of ecological corridors and 'peace parks' to help species adapt to globalization and combat exploitation of the environment by the increasing human population.



The Most Human Human: What Talking with Computers Teaches Us About What It Means to Be Alive

Brian Christian DOUBLEDAY 320 pp. \$27.95 (2011)

When it comes to good conversation, humans are usually better than a machine. Why artificial intelligence so often fails to live up to our emotional expectations, and what that tells us about being human, is the focus of science writer and poet Brian Christian's book. He relates what it was like to take part in a Turing test, which pitted him against a computer in an anonymous chat with a third-party judge in 2009, and the lengths to which he went to make sure that a human won.



Technological Nature: Adaptation and the Future of Human Life

Peter H. Kahn MIT PRESS 240 pp. \$24.95 (2011)

From watching wildlife documentaries to stalking animals with video cameras, we increasingly interact with nature through technology. After studying people's responses to real-time broadcasts of natural scenes and their play with robotic pets, psychologist Peter Kahn concludes that exposure to digital nature is better than having no natural experience at all. But, he finds, you cannot beat the real thing, arguing that we should view this 'technological nature' as a bonus rather than as a substitute for tangible natural wonders.

CORRESPONDENCE

Complexity clouds finance-risk models

We question the usefulness of focusing on sophisticated ecosystem or other models for assessing financial risks (*Nature* **469**, 302–303 and 351–355; 2011) when practical solutions are to hand.

The main bottleneck is more political than technical, driven by a US banking oligarchy that effectively controls the economy. (Europe's and China's banks have more complex interactions with the state.)

We agree that traditional economic models need to be enhanced with interdisciplinary system theory in the medium and long term. But known short-term solutions have demonstrated their value in previous successful policies, and we should relearn and expand some of the old economic wisdom about the specific role of banks.

Banks are important because they create credit in a fractional reserve system, and credit markets are crucial allocators of capital to entrepreneurs. However, this is not considered in macro-economic models used by central banks (indirect influence through interest rates aside). The Austrian school of economists and scholars in the 1930s correctly emphasized that too much credit, encouraged by artificially low interest rates set by the central banks, can lead to bubbles and unsustainable booms. This is what happened in the run-up to the current financial crisis.

One problem that is often overlooked is the misalignment of interests between credit creation by banks for profit maximization versus the amount of credit required by the economy. The Glass–Steagall Act of 1933 reconciled these interests by separating investment, commercial and retail banking and insurance.

Most scholars attribute much of the economic stability after the Second World War to this legislation. Its repeal in 1999 was the culmination of a decade of deregulation justified by the 'great moderation' (go.nature.com/ehkgvv), which turned out to be just a consequence of the actuation of a perpetual money machine that promised unrealistic economic growth.

We need to stop being blinded by complexity so that policy-makers can develop effective responses to the financial crisis and academics can create a genuine science of out-of-equilibrium system economics. **Didier Sornette, Susanne von der Becke** *ETH Zürich, Switzerland.* dsornette@ethz.ch

Science and religion are wise to talk

As recipients of several peer-reviewed grants from the John Templeton Foundation over the past decade, we agree with other recipients who report that the foundation has never sought to interfere with their grant-funded projects (*Nature* **470**, 323–325; 2011).

For those who think that this grant-giving body should not have funded some projects, we might say the same about all the other agencies that have funded our research in biomedical science and geophysics. Receiving a grant does not entail accepting the worth of all the other grants given by the same body.

As far as the mingling of scientific and religious language is concerned, we agree that this is a justifiable concern. In the United Kingdom, the Faraday Institute (our institution) is well known for its criticism of both creationism and intelligent design. Attempts to introduce theological language into the practice of science is as damaging for theology as it is for

science. Each academic discipline has its own specialized language and its own criteria for justifying its claims; mixing them only creates confusion.

However, we disagree with the scientists you cite who oppose any kind of interdisciplinary engagement between science and religion, or who maintain that they are in conflict. Given that almost all organized science education in Europe was carried out by religious institutions for many centuries, and that the premises and practices of science have deep theological roots, such a stance is implausible. The world is as religious as it has ever been — perhaps more so. The scientific community is often embedded in highly religious societies, the United States being a prime example. Friendly dialogue is a wiser strategy than aloof isolationism.

Denis Alexander, Bob White *The Faraday Institute for Science and Religion, Cambridge, UK.* dra24@hermes.cam.ac.uk

Infrastructure vital to genome success

Eric Lander's assessment of the impact of the publication of the human genome sequence (*Nature* **470**, 187–197; 2011) does not comment on the substantial progress made in research-infrastructure development in areas such as resources, technology, computational biology, training, education and ethical, legal and societal issues, as envisaged by the genome community (*Nature* **422**, 835–847; 2004).

Look at the key experimental resource of human cohorts. For example, more than 500,000 volunteers have now signed up for the world's largest prospective study, UK Biobank. Comparable studies are ongoing in Norway and China. A retrospective study,

Biobank Japan, has already proved its investigational value. Researchers are also building a global network of cohort studies through the Canada-based Public Population Project in Genomics.

Biobanking networks are emerging across the European Union. Examples are the Biobanking and Biomolecular Resources Research Infrastructure and the UK DNA Banking Network.

Advanced data-management systems at the Cancer Biomedical Informatics Grid, the US National Center for Biotechnology Information and the European Bioinformatics Institute near Cambridge, UK, are facilitating meta-analysis and experimental design.

Across the European Research Area, under the aegis of the Innovative Medicines Initiative and the European Strategy Forum on Research Infrastructures, an ambitious programme of research-infrastructure development is under way for the entire 'value chain' of medicines. As part of this, biomedical education and training are being revamped through, for example, the public-private European Medicines Research Training Network.

Without such experimentation in and development of its research infrastructure, the fecundity of genome-based research will not fulfil its promise for health and society.

Martin Yuille *University of Manchester, UK.* martin.yuille@manchester.ac.uk
[SEE COMMENT P.159](#)

CONTRIBUTIONS

Correspondence may be submitted to correspondence@nature.com after consulting the author guidelines at <http://go.nature.com/cmchno>.

A light sounding drum

In the effort to demonstrate quantum behaviour in the motion of macroscopic mechanical objects, strong coupling between the objects and an electromagnetic oscillator is advantageous. Such coupling has now been achieved. [SEE LETTER P.204](#)

MILES BLENCOWE

In the first lecture of his 1962 course on gravitation¹, physicist Richard Feynman speculated on the possibility that the laws of quantum mechanics break down for objects heavier than a certain mass scale. Microscopic systems lighter than this mass scale would behave quantum mechanically — for example, by tunnelling through barriers and existing in two states at once. On the other hand, macroscopic mass systems above this scale could behave only in their familiar, classical way as objects localized in space. It would not be possible, even in principle, to prepare states such as that of the notorious Schrödinger cat, consisting of the simultaneous superposition of an alive and a dead feline. Invoking gravity as the enforcer of classical behaviour for macroscopic objects, Feynman combined the fundamental constants of Newton's gravitation, the speed of light and the quantum Planck constant to arrive at a quantum breakdown mass scale (known as the Planck mass) of approximately 0.01 milligrams, or about the weight of a speck of dust.

Whether or not one subscribes to Feynman's speculation, there nevertheless remains the fascinating question of how the microscopic quantum world of subatomic particles, atoms and molecules makes the transition to the macroscopic classical world of everyday objects. On page 204 of this issue, Teufel and colleagues² present experiments involving an electromechanical device that seems ideally suited for exploring this quantum–classical transition.

The device comprises a circular mechanical oscillator rather like a drumhead, 15 micrometres in diameter and 100 nanometres thick (giving it an aspect ratio of 150 to 1), that vibrates at about 11 million cycles per second. The drumhead has a dual purpose, serving not only as a mechanical oscillator but also as an electrical capacitor of the type that consists of two parallel conducting plates separated by a vacuum-filled gap. Electrical contact between this capacitor and another element (an inductor) to form a closed circuit loop produces an electromagnetic oscillator, in which, instead of the oscillating material mass of the drumhead, electrical current flows back and forth,

oscillating at a resonance frequency of about 7 billion cycles per second. This use of a dual-purpose capacitor facilitates a very strong coupling between the two (electrical and mechanical) oscillators, and represents the single most significant advance of the present device over recent related electromechanical schemes^{3,4}. In the latter experiments, the electrical oscillator capacitor and mechanical oscillator capacitor were distinct elements, and, given their disparity in size, the electromechanical coupling was much weaker.

Teufel *et al.*² demonstrate several predicted⁵ consequences of the strong electromechanical coupling that they achieve, together with long-amplitude damping times for the mechanical and superconducting electrical oscillators

The two oscillators are so strongly coupled that it no longer makes sense to think of them as distinct entities, but rather as an electromechanical hybrid.

Probing this strongly driven coupled system, the authors observe two well-separated modes of oscillation that describe electromechanical excitations. The two oscillators are so strongly coupled that it no longer makes sense to think of them as distinct entities, but rather as an electromechanical hybrid.

Although this observed hybridization is well described by classical coupled-oscillator physics, it does establish that the oscillators' mutual coupling strength and their long damping times create favourable conditions for the mechanical system to enter the quantum regime. Possible experiments aimed at accessing the quantum nature of the system include cooling the mechanical drumhead down to its quantum ground state, in which only the quantum 'zero-point' fluctuation energy remains, or the more challenging task of preparing quantum superposition states involving, say, the ground state and the first excited state consisting of a single quantum of vibrational

energy (known as a phonon). In 2010, O'Connell *et al.*⁶ carried out the first such experiments, also using an electromechanical device; their pioneering work was recognized as the 'Breakthrough of the Year' by *Science*⁷. To verify similar quantum behaviour in the present device, one route⁸ might involve hooking up a superconducting 'quantum bit' device to the electrical oscillator circuit; quantum-state preparation and detection of the mechanical oscillator are then achieved by driving and measuring the state of the coupled electrical oscillator.

Given that we now know from experiment⁶ that electromechanical systems of mesoscopic size (ranging from 100 nm to several micrometres) obey the laws of quantum mechanics, is it worth establishing that the same is true for other electromechanical and related optomechanical devices⁹? There are good reasons for continuing to pursue such a goal. In particular, with its long damping times and strong electromechanical coupling, the drumhead device² modified as described above for example, makes it possible not only to verify quantum behaviour, but also — perhaps even more importantly — to watch it rapidly lose its quantum nature, with the mechanical oscillator subsequently behaving in its usual classical manner.

With a mass of around 10^{-13} kilograms, the drumhead is much too light to be used to investigate the possibility considered by Feynman that gravity provides the ultimate quantum–classical divide. However, even well before the Planck mass scale is reached, there are other, less exotic effects that will cause an initially quantum drumhead to become classical, such as the emission of phonon radiation into its supports, or tiny, dissipative defects in the drumhead itself.

Furthermore, with an electrical oscillator to hand, there is no reason that the goal should be to demonstrate only mechanical quantum behaviour. The dynamics of driven, strongly coupled, quantum hybrid systems in the presence of a dissipative environment is even less well understood, and is likely to yield counter-intuitive surprises concerning the quantum–classical divide¹⁰. ■

Miles Blencowe is in the Department of Physics and Astronomy, Dartmouth College,

Hanover, New Hampshire 03755, USA.
e-mail: miles.p.blencowe@dartmouth.edu

1. Feynman, R. P., Morinigo, F. B., Wagner, W. G. & Hatfield, B. *Feynman Lectures on Gravitation* (Westview, 2003).
2. Teufel, J. D. *et al. Nature* **471**, 204–208 (2011).
3. Regal, C. A., Teufel, J. D. & Lehnert, K. W. *Nature Phys.* **4**, 555–560 (2008).

4. Rocheleau, T. *et al. Nature* **463**, 72–75 (2010).
5. Dobrindt, J. M., Wilson-Rae, I. & Kippenberg, T. J. *Phys. Rev. Lett.* **101**, 263602 (2008).
6. O'Connell, A. D. *et al. Nature* **464**, 697–703 (2010).
7. Cho, A. *Science* **330**, 1604 (2010).
8. Armour, A. D. & Blencowe, M. P. *New J. Phys.* **10**, 095004 (2008).
9. Marquardt, F. & Girvin, S. M. *Physics* **2**, 40 (2009).
10. Galve, F., Pachón, L. A. & Zueco, D. *Phys. Rev. Lett.* **105**, 180501 (2010).

IMMUNOLOGY

Context is key in the gut

The vitamin-A metabolite retinoic acid normally favours immune tolerance in the gut. But in coeliac disease — an intestinal inflammatory disorder due to adverse reactivity to a dietary protein — it may do just the opposite. [SEE LETTER P.220](#)

CRAIG L. MAYNARD & CASEY T. WEAVER

The evolution of an adaptive immune system, which can recognize almost any foreign molecule, or antigen, has presented vertebrates with a unique challenge: how to distinguish between harmful and beneficial antigens. The challenge is particularly daunting in the intestine, where antigens derived from the resident microbiota and ingested food probably represent the greatest number and variety of non-self molecules that adaptive immune cells will encounter over a lifetime. In coeliac disease, the immune system of affected individuals fails this challenge, triggering an inflammatory, rather than an anti-inflammatory, response to a dietary component. On page 220 of this issue, DePaolo *et al.*¹ provide evidence that in mice — in a setting similar to coeliac disease — a metabolite that normally contributes to immune balance in the small intestine favours immune pathology instead.

T cells are the immune-system cells that are tasked primarily with promoting tolerance to innocuous foreign antigens, while retaining the capacity to mount immunity to antigens of pathogens that also enter through the intestine. A recent breakthrough in understanding immune tolerance to food antigens (oral tolerance) was the discovery that the vitamin-A metabolite retinoic acid plays a crucial part in this phenomenon^{2,3}.

In this case, retinoic acid, which is continuously produced by intestinal dendritic cells, acts in concert with the cytokine TGF- β , an immune mediator abundant in intestinal tissue. Together they induce the development of regulatory T (T_{reg}) cells from antigen-naïve precursor T cells and so suppress the development of pro-inflammatory effector T cells^{2–4} (Fig. 1). In a subset of genetically susceptible individuals, however, oral tolerance to the common dietary component gluten (a complex of proline- and glutamine-rich

proteins present in wheat, barley and rye) is subverted, resulting in coeliac disease^{5–7}.

Given that retinoic acid normally promotes oral tolerance, it would be reasonable to conjecture that the mechanism leading to coeliac disease involves defects in the production of this metabolite by dendritic cells and/or in the responsiveness of naïve T cells to it in susceptible people. DePaolo *et al.*¹, however, compellingly show in a new mouse model of coeliac disease that neither of these mechanisms is responsible.

Instead, they find that, in the setting of high levels of the cytokine IL-15, which are found in

the intestines of patients with coeliac disease, retinoic acid acts on dendritic cells to promote a destructive immune response to gluten, not tolerance. It seems, therefore, that retinoic acid can either prevent or promote adverse reactions to food antigens, depending on the immunological context.

Although elevated levels of intestinal IL-15 in coeliac disease are well documented^{5,7}, an explanation of how this cytokine might promote immune dysregulation in response to ingested gluten has remained elusive. DePaolo and colleagues found that, in their mice, which were engineered to constitutively express IL-15 in the intestinal immune tissues, orally delivered antigen did not induce the development of T_{reg} cells. Instead, it induced T cells that produce the cytokine IFN- γ — the type of T cells that contribute to the pathology of coeliac disease. Through complementary *ex vivo* and *in vivo* experiments, the authors delineate a mechanism whereby IL-15 supports production of T-cell-derived IFN- γ by directly promoting the expression of the cytokine IL-12 by dendritic cells.

IL-12 is central to the induction and development of T-helper type 1 (T_H1) cells that express IFN- γ , and it can suppress T_{reg} -cell development¹. Remarkably, when DePaolo *et al.* gave the IL-15-expressing mice retinoic acid together with oral antigens, they detected further suppression of antigen-specific T_{reg} cells in favour of increased T_H1 -cell development. This finding correlates with the authors' observation that, in response to both IL-15 and retinoic

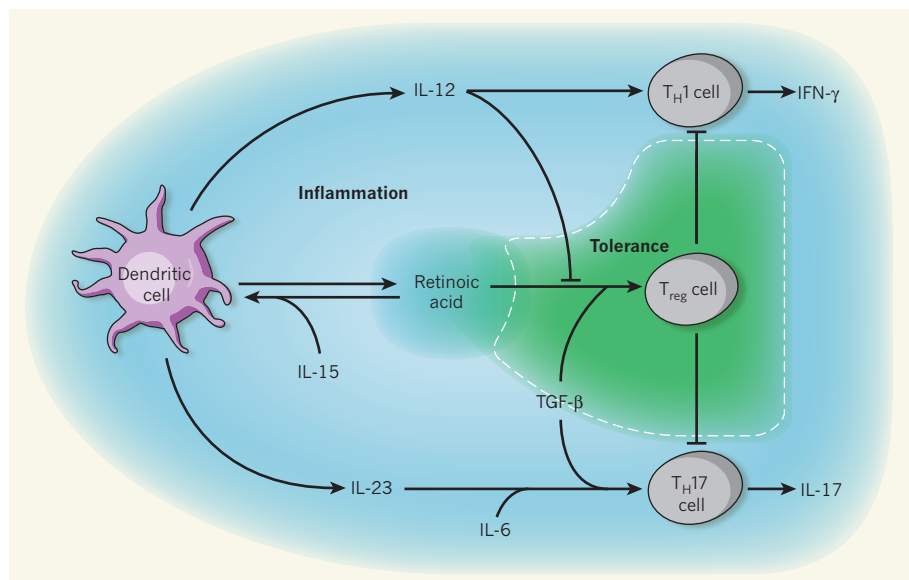


Figure 1 | Dual and competing roles of retinoic acid in mucosal T-cell responses. Together with TGF- β , retinoic acid promotes differentiation of antigen-naïve precursor T cells (not shown) into T_{reg} cells, which help to maintain a state of immune homeostasis by inhibiting the development and function of pro-inflammatory effector T cells. DePaolo *et al.*¹ find, by contrast, that in the presence of high levels of IL-15, retinoic acid supports the induction of the IL-12-family cytokines IL-12 and IL-23, which participate in the induction of pro-inflammatory T_H1 - and T_H17 -cell responses (secretion of IFN- γ and IL-17), respectively. For the induction of oral tolerance, retinoic acid is primarily produced by a subset of mucosal dendritic cells. The cellular source(s) of retinoic acid and IL-15 that facilitate the development of coeliac disease — as well as the specific subset(s) of dendritic cells targeted by the pair — remain to be elucidated.

Hanover, New Hampshire 03755, USA.
e-mail: miles.p.blencowe@dartmouth.edu

1. Feynman, R. P., Morinigo, F. B., Wagner, W. G. & Hatfield, B. *Feynman Lectures on Gravitation* (Westview, 2003).
2. Teufel, J. D. *et al. Nature* **471**, 204–208 (2011).
3. Regal, C. A., Teufel, J. D. & Lehnert, K. W. *Nature Phys.* **4**, 555–560 (2008).

4. Rocheleau, T. *et al. Nature* **463**, 72–75 (2010).
5. Dobrindt, J. M., Wilson-Rae, I. & Kippenberg, T. J. *Phys. Rev. Lett.* **101**, 263602 (2008).
6. O'Connell, A. D. *et al. Nature* **464**, 697–703 (2010).
7. Cho, A. *Science* **330**, 1604 (2010).
8. Armour, A. D. & Blencowe, M. P. *New J. Phys.* **10**, 095004 (2008).
9. Marquardt, F. & Girvin, S. M. *Physics* **2**, 40 (2009).
10. Galve, F., Pachón, L. A. & Zueco, D. *Phys. Rev. Lett.* **105**, 180501 (2010).

IMMUNOLOGY

Context is key in the gut

The vitamin-A metabolite retinoic acid normally favours immune tolerance in the gut. But in coeliac disease — an intestinal inflammatory disorder due to adverse reactivity to a dietary protein — it may do just the opposite. SEE LETTER P.220

CRAIG L. MAYNARD & CASEY T. WEAVER

The evolution of an adaptive immune system, which can recognize almost any foreign molecule, or antigen, has presented vertebrates with a unique challenge: how to distinguish between harmful and beneficial antigens. The challenge is particularly daunting in the intestine, where antigens derived from the resident microbiota and ingested food probably represent the greatest number and variety of non-self molecules that adaptive immune cells will encounter over a lifetime. In coeliac disease, the immune system of affected individuals fails this challenge, triggering an inflammatory, rather than an anti-inflammatory, response to a dietary component. On page 220 of this issue, DePaolo *et al.*¹ provide evidence that in mice — in a setting similar to coeliac disease — a metabolite that normally contributes to immune balance in the small intestine favours immune pathology instead.

T cells are the immune-system cells that are tasked primarily with promoting tolerance to innocuous foreign antigens, while retaining the capacity to mount immunity to antigens of pathogens that also enter through the intestine. A recent breakthrough in understanding immune tolerance to food antigens (oral tolerance) was the discovery that the vitamin-A metabolite retinoic acid plays a crucial part in this phenomenon^{2,3}.

In this case, retinoic acid, which is continuously produced by intestinal dendritic cells, acts in concert with the cytokine TGF- β , an immune mediator abundant in intestinal tissue. Together they induce the development of regulatory T (T_{reg}) cells from antigen-naïve precursor T cells and so suppress the development of pro-inflammatory effector T cells^{2–4} (Fig. 1). In a subset of genetically susceptible individuals, however, oral tolerance to the common dietary component gluten (a complex of proline- and glutamine-rich

proteins present in wheat, barley and rye) is subverted, resulting in coeliac disease^{5–7}.

Given that retinoic acid normally promotes oral tolerance, it would be reasonable to conjecture that the mechanism leading to coeliac disease involves defects in the production of this metabolite by dendritic cells and/or in the responsiveness of naïve T cells to it in susceptible people. DePaolo *et al.*¹, however, compellingly show in a new mouse model of coeliac disease that neither of these mechanisms is responsible.

Instead, they find that, in the setting of high levels of the cytokine IL-15, which are found in

the intestines of patients with coeliac disease, retinoic acid acts on dendritic cells to promote a destructive immune response to gluten, not tolerance. It seems, therefore, that retinoic acid can either prevent or promote adverse reactions to food antigens, depending on the immunological context.

Although elevated levels of intestinal IL-15 in coeliac disease are well documented^{5,7}, an explanation of how this cytokine might promote immune dysregulation in response to ingested gluten has remained elusive. DePaolo and colleagues found that, in their mice, which were engineered to constitutively express IL-15 in the intestinal immune tissues, orally delivered antigen did not induce the development of T_{reg} cells. Instead, it induced T cells that produce the cytokine IFN- γ — the type of T cells that contribute to the pathology of coeliac disease. Through complementary *ex vivo* and *in vivo* experiments, the authors delineate a mechanism whereby IL-15 supports production of T-cell-derived IFN- γ by directly promoting the expression of the cytokine IL-12 by dendritic cells.

IL-12 is central to the induction and development of T-helper type 1 (T_H1) cells that express IFN- γ , and it can suppress T_{reg} -cell development¹. Remarkably, when DePaolo *et al.* gave the IL-15-expressing mice retinoic acid together with oral antigens, they detected further suppression of antigen-specific T_{reg} cells in favour of increased T_H1 -cell development. This finding correlates with the authors' observation that, in response to both IL-15 and retinoic

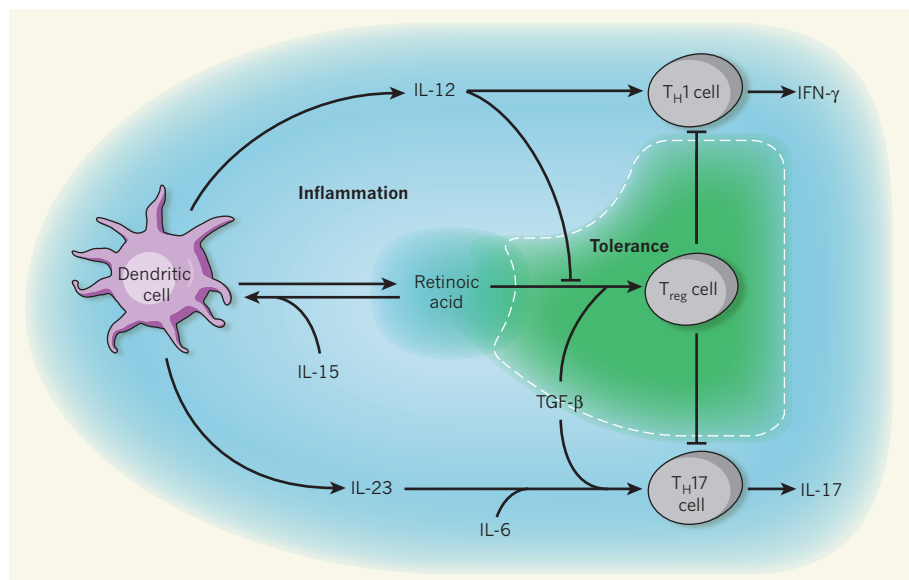


Figure 1 | Dual and competing roles of retinoic acid in mucosal T-cell responses. Together with TGF- β , retinoic acid promotes differentiation of antigen-naïve precursor T cells (not shown) into T_{reg} cells, which help to maintain a state of immune homeostasis by inhibiting the development and function of pro-inflammatory effector T cells. DePaolo *et al.*¹ find, by contrast, that in the presence of high levels of IL-15, retinoic acid supports the induction of the IL-12-family cytokines IL-12 and IL-23, which participate in the induction of pro-inflammatory T_H1 - and T_H17 -cell responses (secretion of IFN- γ and IL-17), respectively. For the induction of oral tolerance, retinoic acid is primarily produced by a subset of mucosal dendritic cells. The cellular source(s) of retinoic acid and IL-15 that facilitate the development of coeliac disease — as well as the specific subset(s) of dendritic cells targeted by the pair — remain to be elucidated.

acid, dendritic cells produce more IL-12. They also found that IL-15 and retinoic acid promote the development of pro-inflammatory T_H17 effector cells by enhancing production of the IL-12-related cytokine IL-23, which, along with TGF- β and IL-6, favours their differentiation. In the presence of IL-15, therefore, retinoic acid seems to act as a co-factor for production of IL-12-family cytokines by dendritic cells, establishing a new and unanticipated function for this metabolite in mucosal immunity (Fig. 1).

A telling aspect of this study emerged when DePaolo and co-workers¹ studied mice that expressed the human MHC class II molecule HLA-DQ8, which can present gluten-derived peptides to mouse T cells. When fed gluten, these animals developed oral tolerance — not immunity. This is consistent with findings^{5–7} in humans that the great majority of individuals who carry the HLA-DQ2 or HLA-DQ8 susceptibility alleles do not develop coeliac disease. Only in gluten-fed mice expressing both HLA-DQ8 and high levels of IL-15 were features of early coeliac disease detected. What's more, administration of retinoic acid to these animals augmented the disease induced by feeding them gluten alone. This finding warns us that treatment of coeliac-disease patients with retinoids could well promote, rather than quell, disease, drawing possible parallels with a report⁸ of patients in whom inflammatory bowel disease developed or was exacerbated following treatment of their acne with retinoids.

Apart from providing insights into the mechanisms underlying coeliac disease, DePaolo and colleagues' work should help to focus future studies on several unanswered questions. What causes the upregulation of IL-15 in patients with coeliac disease? And does the abnormal expression of this cytokine cause or just aggravate the disorder? Consistent with the present study, a gene encoding a subunit of IL-12 has been identified as a risk factor for coeliac disease⁹ whereas, intriguingly, the gene that encodes IL-15 has not. And although the mouse model¹ shows many features of early coeliac disease, it does not exhibit villous atrophy, which is characteristic of more advanced disease in humans and is thought to be mediated by IL-15 activity. So why does it not show atrophy? And what are the implications of this observation for the hypothesis that IL-15 promotes the damage to the intestinal epithelium in coeliac disease?

As for retinoic acid, how exactly does it amplify the pro-inflammatory effects of IL-15 on dendritic cells? DePaolo *et al.* show that retinoic acid enhances activation of the JNK signalling cascade. How does it do this, and is abnormal production of retinoic acid a component of coeliac disease?

Irrespective of the answers to these questions, this study¹ re-emphasizes the fact that the IL-15 pathway is an attractive target for new therapies for coeliac disease. Whereas most

autoimmune diseases are difficult to control because the offending antigen is a constituent of normal tissues and thus cannot be removed, in coeliac disease simple exclusion of dietary gluten is often palliative. Nevertheless, in a large subset of patients, a gluten-free diet does not completely control disease, necessitating other interventions. Abrogation of IL-15 expression or signalling in the intestinal tissues may well stem the adverse immune responses to gluten, as well as providing a means to re-establish oral tolerance in these patients, whether married to efforts to target retinoic acid or not. ■

Craig L. Maynard and Casey T. Weaver
are in the Department of Pathology,
University of Alabama at Birmingham,

Birmingham, Alabama 35294-2170, USA.
e-mails: cmaynard@uab.edu;
cweaver@uab.edu

1. DePaolo, R. R. *et al.* *Nature* **271**, 220–224 (2011).
2. Coombes, J. L. *et al.* *J. Exp. Med.* **204**, 1757–1764 (2007).
3. Sun, C. M. *et al.* *J. Exp. Med.* **204**, 1775–1785 (2007).
4. Mucida, D. *et al.* *Science* **317**, 256–260 (2007).
5. Jabri, B. & Sollid, L. M. *Nature Rev. Immunol.* **9**, 858–870 (2009).
6. Kagnoff, M. F. *J. Clin. Invest.* **117**, 41–49 (2007).
7. Meresse, B., Ripoché, J., Heyman, M. & Cerf-Bennussan, N. *Mucosal Immunol.* **2**, 8–23 (2009).
8. Reddy, D. *et al.* *Am. J. Gastroenterol.* **101**, 1569–1573 (2006).
9. Hunt, K. A. *et al.* *Nature Genet.* **40**, 395–402 (2008).

NEUROSCIENCE

Towards functional connectomics

To understand the brain, the thousands of synaptic connections made by each of billions of neurons should be mapped and related to neuronal function. First steps towards this formidable goal are now reported. SEE ARTICLES P.177 & P.183

H. SEBASTIAN SEUNG

Neurons are classified into cell types, which are traditionally defined by location and shape¹. Rules of synaptic connection that depend on cell type, although important, are not sufficient for understanding neuronal function, because cells of the same type can have diverse functions. In this issue, Bock *et al.*² and Briggman *et al.*³ report an exciting and pioneering approach to finding rules of connection between neurons that depend on their functional properties as well as their cell type.

In the 1970s and 1980s, researchers⁴ painstakingly compiled the 'wiring diagram' — now also known as the connectome⁵ — of the worm *Caenorhabditis elegans*. They imaged extremely thin slices of this organism using serial electron microscopy, virtually reconstructed each neuron by finding all its cross-sections in the images, and found every synaptic connection between neurons.

Serial electron microscopy has been revived and reinvented in recent years, raising hopes of determining the connectomes of mammalian brains⁶. Briggman *et al.* and Bock *et al.* applied improved versions of this method to the mouse retina and primary visual cortex, respectively. They also achieved a new feat — imaging the activity of neurons in a functioning network using two-photon microscopy and calcium-sensitive indicators

before mapping their connectivity. Combining serial electron microscopy with two-photon microscopy is not just a technical tour de force. It is also a way of directly studying the relationship of a neuron's function to its connections — an example of what might be called 'functional connectomics'.

The mammalian retina contains about 50–60 neuronal cell types¹. Most of what we know about the retina's connectivity is formulated as rules governing these cell types. For example, starburst amacrine cells (SACs) make inhibitory synapses onto direction-selective ganglion cells (DSGCs). Both cell types are involved in computing the direction of movement. Each DSGC is preferentially activated by visual stimuli moving in a particular direction, and different DSGCs have different preferred directions. SACs exhibit functional diversity even within a single cell: each of the dendrites of a SAC functions independently, and each responds selectively to motion in a different preferred direction⁷.

Briggman *et al.* (page 183) demonstrate the specificity of SAC–DSGC connections. They show that a DSGC tends to receive many more synapses from a SAC dendrite if their preferred directions are opposite to each other. This supports the hypothesis that DSGCs mostly 'inherit' their direction selectivity from SACs, albeit with an inversion because SACs are inhibitory. By providing this missing piece of evidence, the authors³ have resolved half

acid, dendritic cells produce more IL-12. They also found that IL-15 and retinoic acid promote the development of pro-inflammatory T_H17 effector cells by enhancing production of the IL-12-related cytokine IL-23, which, along with TGF- β and IL-6, favours their differentiation. In the presence of IL-15, therefore, retinoic acid seems to act as a co-factor for production of IL-12-family cytokines by dendritic cells, establishing a new and unanticipated function for this metabolite in mucosal immunity (Fig. 1).

A telling aspect of this study emerged when DePaolo and co-workers¹ studied mice that expressed the human MHC class II molecule HLA-DQ8, which can present gluten-derived peptides to mouse T cells. When fed gluten, these animals developed oral tolerance — not immunity. This is consistent with findings^{5–7} in humans that the great majority of individuals who carry the HLA-DQ2 or HLA-DQ8 susceptibility alleles do not develop coeliac disease. Only in gluten-fed mice expressing both HLA-DQ8 and high levels of IL-15 were features of early coeliac disease detected. What's more, administration of retinoic acid to these animals augmented the disease induced by feeding them gluten alone. This finding warns us that treatment of coeliac-disease patients with retinoids could well promote, rather than quell, disease, drawing possible parallels with a report⁸ of patients in whom inflammatory bowel disease developed or was exacerbated following treatment of their acne with retinoids.

Apart from providing insights into the mechanisms underlying coeliac disease, DePaolo and colleagues' work should help to focus future studies on several unanswered questions. What causes the upregulation of IL-15 in patients with coeliac disease? And does the abnormal expression of this cytokine cause or just aggravate the disorder? Consistent with the present study, a gene encoding a subunit of IL-12 has been identified as a risk factor for coeliac disease⁹ whereas, intriguingly, the gene that encodes IL-15 has not. And although the mouse model¹ shows many features of early coeliac disease, it does not exhibit villous atrophy, which is characteristic of more advanced disease in humans and is thought to be mediated by IL-15 activity. So why does it not show atrophy? And what are the implications of this observation for the hypothesis that IL-15 promotes the damage to the intestinal epithelium in coeliac disease?

As for retinoic acid, how exactly does it amplify the pro-inflammatory effects of IL-15 on dendritic cells? DePaolo *et al.* show that retinoic acid enhances activation of the JNK signalling cascade. How does it do this, and is abnormal production of retinoic acid a component of coeliac disease?

Irrespective of the answers to these questions, this study¹ re-emphasizes the fact that the IL-15 pathway is an attractive target for new therapies for coeliac disease. Whereas most

autoimmune diseases are difficult to control because the offending antigen is a constituent of normal tissues and thus cannot be removed, in coeliac disease simple exclusion of dietary gluten is often palliative. Nevertheless, in a large subset of patients, a gluten-free diet does not completely control disease, necessitating other interventions. Abrogation of IL-15 expression or signalling in the intestinal tissues may well stem the adverse immune responses to gluten, as well as providing a means to re-establish oral tolerance in these patients, whether married to efforts to target retinoic acid or not. ■

Craig L. Maynard and Casey T. Weaver
are in the Department of Pathology,
University of Alabama at Birmingham,

Birmingham, Alabama 35294-2170, USA.
e-mails: cmaynard@uab.edu;
cweaver@uab.edu

1. DePaolo, R. R. *et al.* *Nature* **271**, 220–224 (2011).
2. Coombes, J. L. *et al.* *J. Exp. Med.* **204**, 1757–1764 (2007).
3. Sun, C. M. *et al.* *J. Exp. Med.* **204**, 1775–1785 (2007).
4. Mucida, D. *et al.* *Science* **317**, 256–260 (2007).
5. Jabri, B. & Sollid, L. M. *Nature Rev. Immunol.* **9**, 858–870 (2009).
6. Kagnoff, M. F. *J. Clin. Invest.* **117**, 41–49 (2007).
7. Meresse, B., Ripoché, J., Heyman, M. & Cerf-Bennussan, N. *Mucosal Immunol.* **2**, 8–23 (2009).
8. Reddy, D. *et al.* *Am. J. Gastroenterol.* **101**, 1569–1573 (2006).
9. Hunt, K. A. *et al.* *Nature Genet.* **40**, 395–402 (2008).

NEUROSCIENCE

Towards functional connectomics

To understand the brain, the thousands of synaptic connections made by each of billions of neurons should be mapped and related to neuronal function. First steps towards this formidable goal are now reported. SEE ARTICLES P.177 & P.183

H. SEBASTIAN SEUNG

Neurons are classified into cell types, which are traditionally defined by location and shape¹. Rules of synaptic connection that depend on cell type, although important, are not sufficient for understanding neuronal function, because cells of the same type can have diverse functions. In this issue, Bock *et al.*² and Briggman *et al.*³ report an exciting and pioneering approach to finding rules of connection between neurons that depend on their functional properties as well as their cell type.

In the 1970s and 1980s, researchers⁴ painstakingly compiled the 'wiring diagram' — now also known as the connectome⁵ — of the worm *Caenorhabditis elegans*. They imaged extremely thin slices of this organism using serial electron microscopy, virtually reconstructed each neuron by finding all its cross-sections in the images, and found every synaptic connection between neurons.

Serial electron microscopy has been revived and reinvented in recent years, raising hopes of determining the connectomes of mammalian brains⁶. Briggman *et al.* and Bock *et al.* applied improved versions of this method to the mouse retina and primary visual cortex, respectively. They also achieved a new feat — imaging the activity of neurons in a functioning network using two-photon microscopy and calcium-sensitive indicators

before mapping their connectivity. Combining serial electron microscopy with two-photon microscopy is not just a technical tour de force. It is also a way of directly studying the relationship of a neuron's function to its connections — an example of what might be called 'functional connectomics'.

The mammalian retina contains about 50–60 neuronal cell types¹. Most of what we know about the retina's connectivity is formulated as rules governing these cell types. For example, starburst amacrine cells (SACs) make inhibitory synapses onto direction-selective ganglion cells (DSGCs). Both cell types are involved in computing the direction of movement. Each DSGC is preferentially activated by visual stimuli moving in a particular direction, and different DSGCs have different preferred directions. SACs exhibit functional diversity even within a single cell: each of the dendrites of a SAC functions independently, and each responds selectively to motion in a different preferred direction⁷.

Briggman *et al.* (page 183) demonstrate the specificity of SAC–DSGC connections. They show that a DSGC tends to receive many more synapses from a SAC dendrite if their preferred directions are opposite to each other. This supports the hypothesis that DSGCs mostly 'inherit' their direction selectivity from SACs, albeit with an inversion because SACs are inhibitory. By providing this missing piece of evidence, the authors³ have resolved half

a century of debate over the mechanism of direction selectivity in ganglion cells.

Although a DSGC can receive synapses from SAC dendrites with any preferred direction, the number of synapses tends to be much larger for certain directions. Therefore, methods that merely establish whether a SAC dendrite is connected to a DSGC cannot reveal the specificity of connectivity. It is essential to quantify the strength of interaction by counting the number of synapses involved, as Briggman and co-workers have done.

In the neocortex, these neurons are excitatory and almost all of these are of the pyramidal type. There are also many types of inhibitory neurons and many rules of neocortical connectivity based on cell type⁸. But pyramidal neurons, even in the same cortical location and layer, can differ in their functional properties. For example, a pyramidal neuron in the primary visual cortex is preferentially activated by visual stimuli of one orientation, and the preferred orientations of pyramidal neurons are diverse.

Bock *et al.* (page 177) find that inhibitory neurons receive synapses from pyramidal neurons with a wide range of preferred orientations. This lack of specificity may explain observations^{9–11} that inhibitory neurons are untuned, or only weakly tuned, to stimulus orientation: if an inhibitory neuron indiscriminately sums over synaptic inputs from pyramidal neurons of all preferred orientations, then its output would lack orientation selectivity.

The authors², however, acknowledge their study's limitations. Even if an inhibitory neuron receives synapses from pyramidal neurons with a wide range of preferred orientations, its summed synaptic input could still be biased towards a particular orientation by more or stronger synapses (see Fig. 5d of ref. 2, for example). Other researchers¹² have reported that one type of inhibitory neuron is sharply tuned to a particular orientation. So although Bock *et al.* have not said the last word on the subject, they have made a step in the right direction.

Neural-network theorists are eager to see the controversies resolved, as they have long speculated that inhibitory neurons receive indiscriminate connections from pyramidal neurons and send back inhibition to prevent runaway excitation¹³ or to sharpen response selectivity^{14,15}. If this idea is correct, inhibitory neurons have a supporting role in visual computations: they are not primarily responsible for generating selectivity to visual features, but rather they help pyramidal neurons to achieve it.

The two teams^{2,3} imaged relatively small volumes of the brain (less than 1% of a cubic millimetre). Furthermore, the volumes were relatively thin — about 50–60 μm — along one dimension. This is not a problem for the retina, which is a thin sheet. (The inner plexiform layer of the retina, which contains the SAC–DSGC connections, is even thinner.) But many

cortical axons left the confines of the volume that Bock *et al.* imaged. This limited the size of their sample of connections, and biased it towards nearby neurons.

Why not study larger volumes? Although the present studies benefited from recent inventions, they still required heroic efforts to acquire and analyse terabytes of image data. To tackle larger volumes, the speed of both image acquisition and analysis must be increased. Analysis should be accelerated by augmenting human intelligence — as used in these papers — with artificial intelligence, and progress is being made along these lines¹⁶.

Another limitation of the studies^{2,3} is in the fraction of connections mapped within the volume. Mapping a larger fraction would allow the application of sophisticated computational methods for analysing connectivity to understand function¹⁷.

Briggman *et al.* show that DSGCs inherit their direction selectivity from SAC dendrites, whereas Bock *et al.* find that inhibitory neurons in the visual cortex squander their inheritance, discarding the orientation selectivity of their inputs. To understand vision, it will be essential to investigate whether and how connectivity enables a neuron to compute a property that is not already present in any single one of its inputs. Iteration of such connectivity could yield neurons that are selective to more and more complex features, as in many neural-network models of the visual system¹⁸. It is hoped that functional connectomics will finally succeed in revealing how this happens.

These papers^{2,3} have introduced a general approach to relating the structure of neural networks to their function: search for rules of connectivity that depend on functional properties of neurons. Finding such rules will be more arduous than finding connections between brain regions, or rules of connection between neuronal cell types. But it is crucial for testing the claim that “Nothing defines the function of a neuron more than its connections with other neurons”¹⁹. This battle cry will be heard more often as the nascent field of connectomics matures. ■

H. Sebastian Seung is in the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
e-mail: seung@mit.edu

1. Masland, R. H. *Curr. Biol.* **14**, R497 (2004).
2. Bock, D. D. *et al. Nature* **271**, 177–182 (2011).
3. Briggman, K. L., Helmstaedter, M. & Denk, W. *Nature* **271**, 183–188 (2011).
4. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. *Phil. Trans. R. Soc. Lond. B* **314**, 1–340 (1986).
5. Sporns, O., Tononi, G. & Kotter, R. *PLoS Comput. Biol.* **1**, e42 (2005).
6. Helmstaedter, M., Briggman, K. L. & Denk, W. *Curr. Opin. Neurobiol.* **18**, 633–641 (2008).
7. Euler, T., Detwiler, P. B. & Denk, W. *Nature* **418**, 845–852 (2002).
8. Douglas, R. J. & Martin, K. A. C. *Annu. Rev.*



50 Years Ago

The Mango. By Dr. Lal Behari Singh — It is pleasing, as it is unexpected, to find a new series of books on world crops starting off not with such solid fare as wheat or potatoes or cabbage but with something as exotic and appetizing as the mango. This “choicest fruit of Hindoosthan” has spread far beyond the bounds of India and has become one of the most cherished fruits of tropical lands. It is still found in greatest variety and excellence in India, and this book comes, fitly enough, from the pen of a distinguished Indian horticulturist. This book may be counted on to commend itself as the most complete study of the mango so far published ... While the botanical chapters refer mainly to the common mango, *M. indica*, the author directs attention to fifteen other species of *Mangifera* with edible fruits. Evidently there is great wealth of related material here, of possible value for future breeding ... The section on utilization is necessarily brief, since the best way to use a mango is still to eat it as promptly and as dexterously as possible, but some mango recipes are also given ... This book deserves to be welcomed and to be gratefully added to the small, but growing, collection of hand-books on the tropical crops.

From *Nature* 11 March 1961

100 Years Ago

Can any correspondent of *Nature* recall a case of a cat playing with a shadow? I know of a cat — a blue Persian — which appears to wait until the morning sun throws the shadow of a cage-bird on the wall of a room, and then seems to play at catching the shadow of the bird as it moves about.

From *Nature* 9 March 1911

- Neurosci.* **27**, 419–451 (2004).
9. Sohya, K., Kameyama, K., Yanagawa, Y., Obata, K. & Tsumoto, T. *J. Neurosci.* **27**, 2145–2149 (2007).
 10. Niell, C. M. & Stryker, M. P. *J. Neurosci.* **28**, 7520–7536 (2008).
 11. Kerlin, A. M., Andermann, M. L., Berezovskii, V. K. & Reid, R. C. *Neuron* **67**, 858–871 (2010).
 12. Runyan, C. A. *et al.* *Neuron* **67**, 847–857 (2010).
 13. Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. & Suarez, H. H. *Science* **269**, 981–985 (1995).
 14. Somers, D. C., Nelson, S. B. & Sur, M. J. *Neurosci.* **15**, 5448 (1995).
 15. Ben-Yishai, R., Bar-Or, R. L. & Sompolinsky, H. *Proc. Natl Acad. Sci. USA* **92**, 3844–3848 (1995).
 16. Jain, V., Seung, H. S. & Turaga, S. C. *Curr. Opin. Neurobiol.* **20**, 653–666 (2010).
 17. Seung, H. S. *Neuron* **62**, 17–29 (2009).
 18. Fukushima, K. *Biol. Cybernet.* **36**, 193–202 (1980).
 19. Mesulam, M. *Ann. Neurol.* **57**, 5–7 (2005).

MATERIALS SCIENCE

Bubble wrap of cell-like aggregates

Using a microfluidic device, tiny polymeric capsules have been made in which different compounds can be isolated in separate, membrane-bound compartments — a prerequisite for the development of artificial cell aggregates.

TAKAMASA HARADA & DENNIS E. DISCHER

Adding milk to coffee or tea leads to a homogeneous (and tasty) mixture, but how could one keep such fluids apart when they are combined? Aqueous fluids tend to flow and mix together even on the micro-metre scale of biological cells, and so keeping such fluids apart requires a robust partition. Cells in tissues achieve this by using flexible membranes not only to define and delimit each cell, but also to compartmentalize their nuclei and other organelles. Reporting in *Angewandte Chemie*, Shum *et al.*¹ describe similar hierarchical assemblies of synthetic cell-like structures, which use polymeric

surfactants as the building blocks of membranes instead of natural lipids and proteins. These robust, deformable microcapsules adhere tightly to each other while confining distinct aqueous fluids, in the same way that cells cluster together in tissues while enclosing separate portions of cytoplasm.

To achieve this feat, Shum *et al.* used a microfluidic device to make water-in-oil-in-water (W/O/W) double emulsions as templates for the polymeric assemblies (Fig. 1). Such double emulsions consist of a water droplet (or droplets) surrounded by an oil phase, which is itself suspended in a second aqueous solution. The authors precisely controlled the initial size of their emulsions by changing the

volume fraction of each phase being passed through the glass capillaries of the microfluidic device, and/or by changing the capillary diameter and flow rate. Importantly, the number of water droplets incorporated into the double emulsions could also be varied by controlling the flow rate of the phases.

The real key to their success, however, was in the choice of ingredients for the oil phase. The authors used two organic solvents mixed with a block copolymer (a polymer in which two or more chemically distinct polymer chains are connected together). The solubility of the copolymer was greater in one of the solvents (chloroform) than in the other (hexane). Because the copolymer was amphiphilic — consisting of a hydrophobic chain covalently linked to a hydrophilic chain — it became concentrated at the oil–water interfaces in the W/O/W double emulsions.

Once the double emulsions had formed, the volatile chloroform started to evaporate. Shum *et al.*¹ observed that, as the chloroform left the system, the inner water droplets became coated with a dense monolayer (leaflet) of copolymer. Because the solubility of the copolymer in the remaining hexane was lower than in the original chloroform–hexane mixture, the water droplets started to stick together. The droplets also adhered to the interface of the oil phase with the surrounding water, where another leaflet of copolymer had formed.

The authors then eliminated the remaining hexane either by evaporation or by using shear in the microfluidic flow of their system. The end result was a highly cohesive assembly of polymer vesicles (polymersomes), wrapped together in a shared outer copolymer leaflet. The orientations of the inner droplets to each other and the contact angles of the interfaces between the droplets determined the eventual shapes of the polymersomes in the aggregates. This meant that the authors could make polymersomes of different shapes, but with the same number of compartments.

Shum and colleagues' ensembles resemble aggregates of soap bubbles and also adherent cells such as those seen in the eye². The comparison with adherent cells is telling, because the formation of structured subsets of cells in tissues is a complex process that is also controlled to some extent by the physics of adhesion. Networks of biochemical reactions direct cell crawling and other key processes, but developmental biologists have also noted that simple mixtures of cells sort themselves in a manner reminiscent of the behaviour of immiscible liquids in emulsions. It has been thought for decades that cell sorting *in vivo* might occur because cells preferentially adopt configurations that minimize the surface and bulk mechanical energies of cell clusters. Evidence of this has come most recently from *in vitro* studies³ in which the adhesion strengths of cells were altered in systems containing small numbers of different cell types.

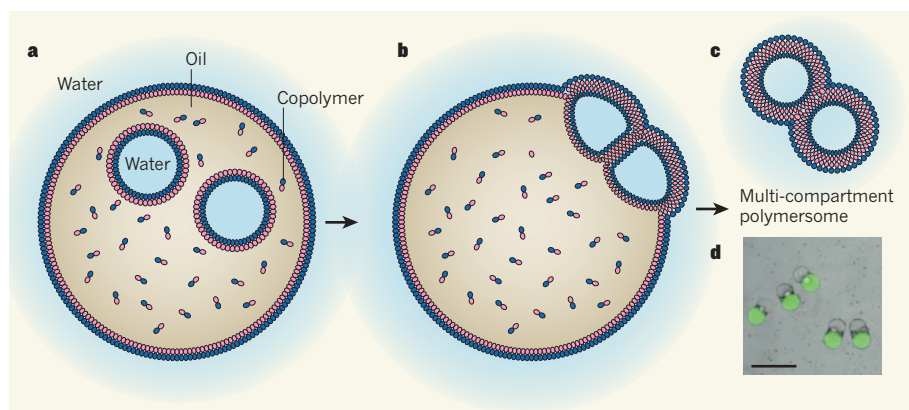


Figure 1 | The synthesis of polymeric cell-like aggregates. Shum *et al.*¹ have used a microfluidic device to prepare multi-compartment polymer vesicles. **a**, The authors began by making water-in-oil-in-water (W/O/W) double emulsions — water droplets suspended in an oil phase, which is itself suspended in water. The oil phase consisted of two solvents and a copolymer, which concentrates at the interfaces of the emulsions as monolayers. **b**, As the more volatile solvent in the oil phase evaporates, the monolayers become adhesive and stick to each other. **c**, The authors then removed the remaining solvent from the oil phase, generating multi-compartment vesicles called polymersomes, whose membranes consist of bilayers of copolymer molecules. (Graphics in **a**–**c** adapted from Scheme 1 of ref. 1.) **d**, This picture overlays optical and fluorescent images of polymersomes. One compartment contains a fluorescent solute, with the other containing a non-fluorescent solute. No cross-contamination occurs. Scale bar, 200 μm . (Image from ref. 1.)

a century of debate over the mechanism of direction selectivity in ganglion cells.

Although a DSGC can receive synapses from SAC dendrites with any preferred direction, the number of synapses tends to be much larger for certain directions. Therefore, methods that merely establish whether a SAC dendrite is connected to a DSGC cannot reveal the specificity of connectivity. It is essential to quantify the strength of interaction by counting the number of synapses involved, as Briggman and co-workers have done.

In the neocortex, these neurons are excitatory and almost all of these are of the pyramidal type. There are also many types of inhibitory neurons and many rules of neocortical connectivity based on cell type⁸. But pyramidal neurons, even in the same cortical location and layer, can differ in their functional properties. For example, a pyramidal neuron in the primary visual cortex is preferentially activated by visual stimuli of one orientation, and the preferred orientations of pyramidal neurons are diverse.

Bock *et al.* (page 177) find that inhibitory neurons receive synapses from pyramidal neurons with a wide range of preferred orientations. This lack of specificity may explain observations^{9–11} that inhibitory neurons are untuned, or only weakly tuned, to stimulus orientation: if an inhibitory neuron indiscriminately sums over synaptic inputs from pyramidal neurons of all preferred orientations, then its output would lack orientation selectivity.

The authors², however, acknowledge their study's limitations. Even if an inhibitory neuron receives synapses from pyramidal neurons with a wide range of preferred orientations, its summed synaptic input could still be biased towards a particular orientation by more or stronger synapses (see Fig. 5d of ref. 2, for example). Other researchers¹² have reported that one type of inhibitory neuron is sharply tuned to a particular orientation. So although Bock *et al.* have not said the last word on the subject, they have made a step in the right direction.

Neural-network theorists are eager to see the controversies resolved, as they have long speculated that inhibitory neurons receive indiscriminate connections from pyramidal neurons and send back inhibition to prevent runaway excitation¹³ or to sharpen response selectivity^{14,15}. If this idea is correct, inhibitory neurons have a supporting role in visual computations: they are not primarily responsible for generating selectivity to visual features, but rather they help pyramidal neurons to achieve it.

The two teams^{2,3} imaged relatively small volumes of the brain (less than 1% of a cubic millimetre). Furthermore, the volumes were relatively thin — about 50–60 μm — along one dimension. This is not a problem for the retina, which is a thin sheet. (The inner plexiform layer of the retina, which contains the SAC–DSGC connections, is even thinner.) But many

cortical axons left the confines of the volume that Bock *et al.* imaged. This limited the size of their sample of connections, and biased it towards nearby neurons.

Why not study larger volumes? Although the present studies benefited from recent inventions, they still required heroic efforts to acquire and analyse terabytes of image data. To tackle larger volumes, the speed of both image acquisition and analysis must be increased. Analysis should be accelerated by augmenting human intelligence — as used in these papers — with artificial intelligence, and progress is being made along these lines¹⁶.

Another limitation of the studies^{2,3} is in the fraction of connections mapped within the volume. Mapping a larger fraction would allow the application of sophisticated computational methods for analysing connectivity to understand function¹⁷.

Briggman *et al.* show that DSGCs inherit their direction selectivity from SAC dendrites, whereas Bock *et al.* find that inhibitory neurons in the visual cortex squander their inheritance, discarding the orientation selectivity of their inputs. To understand vision, it will be essential to investigate whether and how connectivity enables a neuron to compute a property that is not already present in any single one of its inputs. Iteration of such connectivity could yield neurons that are selective to more and more complex features, as in many neural-network models of the visual system¹⁸. It is hoped that functional connectomics will finally succeed in revealing how this happens.

These papers^{2,3} have introduced a general approach to relating the structure of neural networks to their function: search for rules of connectivity that depend on functional properties of neurons. Finding such rules will be more arduous than finding connections between brain regions, or rules of connection between neuronal cell types. But it is crucial for testing the claim that “Nothing defines the function of a neuron more than its connections with other neurons”¹⁹. This battle cry will be heard more often as the nascent field of connectomics matures. ■

H. Sebastian Seung is in the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
e-mail: seung@mit.edu

1. Masland, R. H. *Curr. Biol.* **14**, R497 (2004).
2. Bock, D. D. *et al.* *Nature* **271**, 177–182 (2011).
3. Briggman, K. L., Helmstaedter, M. & Denk, W. *Nature* **271**, 183–188 (2011).
4. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. *Phil. Trans. R. Soc. Lond. B* **314**, 1–340 (1986).
5. Sporns, O., Tononi, G. & Kotter, R. *PLoS Comput. Biol.* **1**, e42 (2005).
6. Helmstaedter, M., Briggman, K. L. & Denk, W. *Curr. Opin. Neurobiol.* **18**, 633–641 (2008).
7. Euler, T., Detwiler, P. B. & Denk, W. *Nature* **418**, 845–852 (2002).
8. Douglas, R. J. & Martin, K. A. C. *Annu. Rev.*



50 Years Ago

The Mango. By Dr. Lal Behari Singh — It is pleasing, as it is unexpected, to find a new series of books on world crops starting off not with such solid fare as wheat or potatoes or cabbage but with something as exotic and appetizing as the mango. This “choicest fruit of Hindoosthan” has spread far beyond the bounds of India and has become one of the most cherished fruits of tropical lands. It is still found in greatest variety and excellence in India, and this book comes, fitly enough, from the pen of a distinguished Indian horticulturist. This book may be counted on to commend itself as the most complete study of the mango so far published ... While the botanical chapters refer mainly to the common mango, *M. indica*, the author directs attention to fifteen other species of *Mangifera* with edible fruits. Evidently there is great wealth of related material here, of possible value for future breeding ... The section on utilization is necessarily brief, since the best way to use a mango is still to eat it as promptly and as dexterously as possible, but some mango recipes are also given ... This book deserves to be welcomed and to be gratefully added to the small, but growing, collection of hand-books on the tropical crops.

From *Nature* 11 March 1961

100 Years Ago

Can any correspondent of *Nature* recall a case of a cat playing with a shadow? I know of a cat — a blue Persian — which appears to wait until the morning sun throws the shadow of a cage-bird on the wall of a room, and then seems to play at catching the shadow of the bird as it moves about.

From *Nature* 9 March 1911

- Neurosci.* **27**, 419–451 (2004).
 9. Sohya, K., Kameyama, K., Yanagawa, Y., Obata, K. & Tsumoto, T. *J. Neurosci.* **27**, 2145–2149 (2007).
 10. Niell, C. M. & Stryker, M. P. *J. Neurosci.* **28**, 7520–7536 (2008).
 11. Kerlin, A. M., Andermann, M. L., Berezovskii, V. K. & Reid, R. C. *Neuron* **67**, 858–871 (2010).
 12. Runyan, C. A. *et al.* *Neuron* **67**, 847–857 (2010).
 13. Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. & Suarez, H. H. *Science* **269**, 981–985 (1995).
 14. Somers, D. C., Nelson, S. B. & Sur, M. J. *Neurosci.* **15**, 5448 (1995).
 15. Ben-Yishai, R., Bar-Or, R. L. & Sompolinsky, H. *Proc. Natl Acad. Sci. USA* **92**, 3844–3848 (1995).
 16. Jain, V., Seung, H. S. & Turaga, S. C. *Curr. Opin. Neurobiol.* **20**, 653–666 (2010).
 17. Seung, H. S. *Neuron* **62**, 17–29 (2009).
 18. Fukushima, K. *Biol. Cybernet.* **36**, 193–202 (1980).
 19. Mesulam, M. *Ann. Neurol.* **57**, 5–7 (2005).

MATERIALS SCIENCE

Bubble wrap of cell-like aggregates

Using a microfluidic device, tiny polymeric capsules have been made in which different compounds can be isolated in separate, membrane-bound compartments — a prerequisite for the development of artificial cell aggregates.

TAKAMASA HARADA & DENNIS E. DISCHER

Adding milk to coffee or tea leads to a homogeneous (and tasty) mixture, but how could one keep such fluids apart when they are combined? Aqueous fluids tend to flow and mix together even on the micro-metre scale of biological cells, and so keeping such fluids apart requires a robust partition. Cells in tissues achieve this by using flexible membranes not only to define and delimit each cell, but also to compartmentalize their nuclei and other organelles. Reporting in *Angewandte Chemie*, Shum *et al.*¹ describe similar hierarchical assemblies of synthetic cell-like structures, which use polymeric

surfactants as the building blocks of membranes instead of natural lipids and proteins. These robust, deformable microcapsules adhere tightly to each other while confining distinct aqueous fluids, in the same way that cells cluster together in tissues while enclosing separate portions of cytoplasm.

To achieve this feat, Shum *et al.* used a microfluidic device to make water-in-oil-in-water (W/O/W) double emulsions as templates for the polymeric assemblies (Fig. 1). Such double emulsions consist of a water droplet (or droplets) surrounded by an oil phase, which is itself suspended in a second aqueous solution. The authors precisely controlled the initial size of their emulsions by changing the

volume fraction of each phase being passed through the glass capillaries of the microfluidic device, and/or by changing the capillary diameter and flow rate. Importantly, the number of water droplets incorporated into the double emulsions could also be varied by controlling the flow rate of the phases.

The real key to their success, however, was in the choice of ingredients for the oil phase. The authors used two organic solvents mixed with a block copolymer (a polymer in which two or more chemically distinct polymer chains are connected together). The solubility of the copolymer was greater in one of the solvents (chloroform) than in the other (hexane). Because the copolymer was amphiphilic — consisting of a hydrophobic chain covalently linked to a hydrophilic chain — it became concentrated at the oil–water interfaces in the W/O/W double emulsions.

Once the double emulsions had formed, the volatile chloroform started to evaporate. Shum *et al.*¹ observed that, as the chloroform left the system, the inner water droplets became coated with a dense monolayer (leaflet) of copolymer. Because the solubility of the copolymer in the remaining hexane was lower than in the original chloroform–hexane mixture, the water droplets started to stick together. The droplets also adhered to the interface of the oil phase with the surrounding water, where another leaflet of copolymer had formed.

The authors then eliminated the remaining hexane either by evaporation or by using shear in the microfluidic flow of their system. The end result was a highly cohesive assembly of polymer vesicles (polymersomes), wrapped together in a shared outer copolymer leaflet. The orientations of the inner droplets to each other and the contact angles of the interfaces between the droplets determined the eventual shapes of the polymersomes in the aggregates. This meant that the authors could make polymersomes of different shapes, but with the same number of compartments.

Shum and colleagues' ensembles resemble aggregates of soap bubbles and also adherent cells such as those seen in the eye². The comparison with adherent cells is telling, because the formation of structured subsets of cells in tissues is a complex process that is also controlled to some extent by the physics of adhesion. Networks of biochemical reactions direct cell crawling and other key processes, but developmental biologists have also noted that simple mixtures of cells sort themselves in a manner reminiscent of the behaviour of immiscible liquids in emulsions. It has been thought for decades that cell sorting *in vivo* might occur because cells preferentially adopt configurations that minimize the surface and bulk mechanical energies of cell clusters. Evidence of this has come most recently from *in vitro* studies³ in which the adhesion strengths of cells were altered in systems containing small numbers of different cell types.

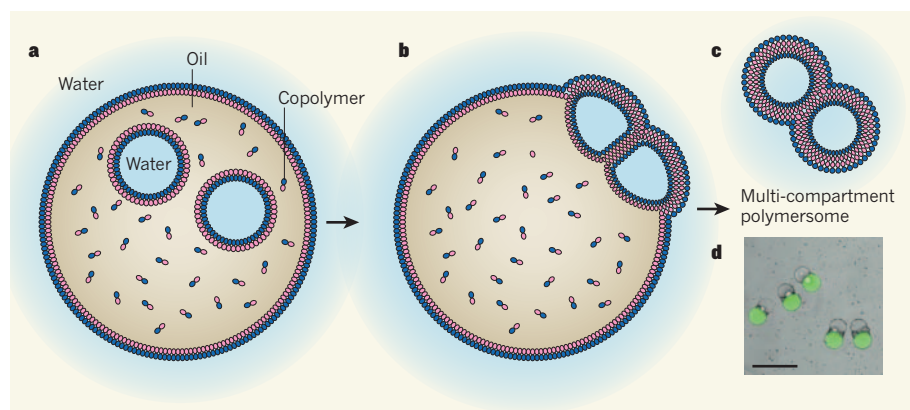


Figure 1 | The synthesis of polymeric cell-like aggregates. Shum *et al.*¹ have used a microfluidic device to prepare multi-compartment polymer vesicles. **a**, The authors began by making water-in-oil-in-water (W/O/W) double emulsions — water droplets suspended in an oil phase, which is itself suspended in water. The oil phase consisted of two solvents and a copolymer, which concentrates at the interfaces of the emulsions as monolayers. **b**, As the more volatile solvent in the oil phase evaporates, the monolayers become adhesive and stick to each other. **c**, The authors then removed the remaining solvent from the oil phase, generating multi-compartment vesicles called polymersomes, whose membranes consist of bilayers of copolymer molecules. (Graphics in **a**–**c** adapted from Scheme 1 of ref. 1.) **d**, This picture overlays optical and fluorescent images of polymersomes. One compartment contains a fluorescent solute, with the other containing a non-fluorescent solute. No cross-contamination occurs. Scale bar, 200 μm . (Image from ref. 1.)

The results of these studies indicated that cells sort themselves according to the surface features of their aggregates.

Inspired by cells, materials scientists have spent more than a decade making microcapsules in which self-assembled polymer bilayers surround a single compartment. Such polymersomes are now made from a wide range of amphiphilic polymers, using methods⁴ that range from the hydration of dried films of a copolymer to solvent extraction from a polymer-solubilizing mixture of water and solvent. Polymersomes are generally more robust than lipid vesicles, and exhibit a wide range of properties that are governed by both the chemical properties and molecular masses of the polymer from which they are made. What's more, polymersomes have been constructed⁵ in which calcium or copper metal ions induce the formation of domains similar to the rafts hypothesized to exist in cell membranes.

One reason for the broad interest in polymersomes is that they might eventually be used to deliver drugs to target tissues in humans. Shum and colleagues¹ also have this goal in mind, because they used a biocompatible copolymer (PEG-*b*-PLA) that has been approved by the US Food and Drug Administration for drug-delivery purposes. Indeed, PEG-*b*-PLA has previously been used to make polymersomes that have delivered therapeutic nucleic acids, such as interfering RNA and antisense oligonucleotides, to cells in animals⁶.

Many other applications of polymersomes are also being developed, including synthetic cells, but Shum and colleagues' vesicle aggregates¹ offer intriguing new possibilities: unlike conventional vesicles, they could encapsulate several different cargoes without cross-contamination. The authors hypothesize that their vesicle aggregates could serve as micrometre-scale reactors, for example, in which different reactants are loaded into separate compartments, the number of which sets the stoichiometry of the reaction. The loaded polymersomes could be delivered to a target site, whereupon controlled degradation of the polymer triggers a reaction as the encapsulated chemicals are released.

To prove that they could indeed load different compounds into different compartments of the same polymersome, Shum *et al.* constructed a microfluidic device that uses separate microchannels to inject water droplets containing different solutes — one of which was fluorescent — into the oil phase of double emulsions. The authors made polymersomes from the emulsions as before, and then used optical and fluorescent microscopy to prove that the fluorescent cargo was trapped exclusively in one compartment.

A grander challenge for vesicle assembly is to prepare functional mimics of cell aggregates in tissues, in which each vesicle performs

a specialized function. Such differentiation of cells in close proximity is what makes tissues so complex. Shum and colleagues' microfluidic method for assembling polymersome aggregates under mild reaction conditions is likely to stimulate much research in this area, and to be used for many other applications, for years to come. ■

Takamasa Harada and Dennis E. Discher
are in the Department of Chemical and Biomolecular Engineering, University of

Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

e-mail: discher@seas.upenn.edu

1. Shum, H. C., Zhao, Y., Kim, S.-H. & Weitz, D. A. *Angew. Chem. Int. Edn* **50**, 1648–1651 (2011).
2. Hayashi, T. & Carthew, R. W. *Nature* **431**, 647–652 (2004).
3. Manning, M. L. *et al. Proc. Natl Acad. Sci. USA* **107**, 12517–12522 (2010).
4. Discher, D. E. & Eisenberg, A. *Science* **297**, 967–973 (2002).
5. Christian, D. A. *et al. Nature Mater.* **8**, 843–849 (2009).
6. Kim, Y. *et al. J. Controlled Release* **134**, 132–140 (2009).

HOST-PATHOGEN INTERACTION

Culprit within a culprit

The parasitic infection mucocutaneous leishmaniasis can vary in severity. It emerges that the levels of an RNA virus within the parasite affect both the host's immune response and the parasite's persistence.

MARTIN OLIVIER

Protozoan parasites of the genus *Leishmania* are responsible for the infectious disease leishmaniasis. The disease mainly affects those living in tropical and sub-tropical parts of the world, and a unique form of it called mucocutaneous leishmaniasis causes a disfiguring disease in the Americas. Despite the century-old knowledge that *Leishmania* parasites of the *Viannia* subgenus cause leishmaniasis, how the disease progresses remains enigmatic. How does a primary cutaneous lesion lead to a secondary, metastasizing form reappearing months later in the nasopharyngeal tissues? Writing in *Science*, Ives *et al.*¹ shed light on this process, showing that the parasite recruits help from within to counteract the host's immune responses, thus enabling it to spread farther afield.

Mucocutaneous leishmaniasis is associated with a persistent inflammatory response characterized by increased expression of pro-inflammatory mediators (TNF- α and some chemokines)^{2,3} that play a pivotal part in recruiting immune-system cells such as macrophages to the site of infection. Using previous approaches^{4,5}, Ives *et al.* isolated clones of *Leishmania guyanensis* — a parasite of the *Viannia* subgenus — from infected hamsters.

The clones were either non-metastatic (*L.g.M*–) or metastatic (*L.g.M*+), and Ives *et al.* found that the *L.g.M*+ clones harboured a virus called *Leishmania* RNA virus 1 (LRV1). When the authors infected mouse macrophages with these clones, the cells rapidly expressed several cytokines and chemokines (including TNF- α and CXCL10) that are relevant to the pathology of mucocutaneous leishmaniasis. The strongest response was to *L.g.M*+ clones and to another clone isolated from patients

with mucocutaneous leishmaniasis. By contrast, *L.g.M*– clones and their corresponding human isolates from cutaneous lesions led to only a modest, albeit significant, response compared with uninfected cells; this is consistent with earlier findings comparing different *Leishmania* species⁶.

Ives *et al.*¹ show that the ability of the *L.g.M*+ clones to induce increased expression of pro-inflammatory molecules depends on their internalization into macrophages and their sequestration within phagolysosomes. A phagolysosome is an organelle in which engulfed pathogens can be destroyed; *Leishmania*, however, has evolved to survive its harsh environment.

An inflammatory immune response is often triggered by the recognition of a pathogen by specific receptors of the innate immune system. One such receptor family comprises the Toll-like receptors (TLRs), which are found on the cell surface or on vesicle membranes of macrophages and other cells of innate immunity. To determine how the inflammatory response to *L. guyanensis* comes about, Ives and colleagues studied macrophages that were functionally deficient in the various phagosomal TLRs (TLR3, 7 and 9) or in intracellular adaptors related to TLR signalling (MyD88 and TRIF).

They found that TLR3 and TRIF are essential for maximal production not just of pro-inflammatory mediators induced by *L.g.M*+, but also of IFN- β — a cytokine produced in response to TLR3 activation and which potentially can cause organ damage. Intriguingly, previous work on virus-free *Leishmania* species suggested that IFN- β ⁷ and TLR3⁸ could be involved, respectively, in the progression of leishmaniasis and in the parasite's recognition by IFN- γ -primed macrophages. Those studies

The results of these studies indicated that cells sort themselves according to the surface features of their aggregates.

Inspired by cells, materials scientists have spent more than a decade making microcapsules in which self-assembled polymer bilayers surround a single compartment. Such polymersomes are now made from a wide range of amphiphilic polymers, using methods⁴ that range from the hydration of dried films of a copolymer to solvent extraction from a polymer-solubilizing mixture of water and solvent. Polymersomes are generally more robust than lipid vesicles, and exhibit a wide range of properties that are governed by both the chemical properties and molecular masses of the polymer from which they are made. What's more, polymersomes have been constructed⁵ in which calcium or copper metal ions induce the formation of domains similar to the rafts hypothesized to exist in cell membranes.

One reason for the broad interest in polymersomes is that they might eventually be used to deliver drugs to target tissues in humans. Shum and colleagues¹ also have this goal in mind, because they used a biocompatible copolymer (PEG-*b*-PLA) that has been approved by the US Food and Drug Administration for drug-delivery purposes. Indeed, PEG-*b*-PLA has previously been used to make polymersomes that have delivered therapeutic nucleic acids, such as interfering RNA and antisense oligonucleotides, to cells in animals⁶.

Many other applications of polymersomes are also being developed, including synthetic cells, but Shum and colleagues' vesicle aggregates¹ offer intriguing new possibilities: unlike conventional vesicles, they could encapsulate several different cargoes without cross-contamination. The authors hypothesize that their vesicle aggregates could serve as micrometre-scale reactors, for example, in which different reactants are loaded into separate compartments, the number of which sets the stoichiometry of the reaction. The loaded polymersomes could be delivered to a target site, whereupon controlled degradation of the polymer triggers a reaction as the encapsulated chemicals are released.

To prove that they could indeed load different compounds into different compartments of the same polymersome, Shum *et al.* constructed a microfluidic device that uses separate microchannels to inject water droplets containing different solutes — one of which was fluorescent — into the oil phase of double emulsions. The authors made polymersomes from the emulsions as before, and then used optical and fluorescent microscopy to prove that the fluorescent cargo was trapped exclusively in one compartment.

A grander challenge for vesicle assembly is to prepare functional mimics of cell aggregates in tissues, in which each vesicle performs

a specialized function. Such differentiation of cells in close proximity is what makes tissues so complex. Shum and colleagues' microfluidic method for assembling polymersome aggregates under mild reaction conditions is likely to stimulate much research in this area, and to be used for many other applications, for years to come. ■

Takamasa Harada and Dennis E. Discher
are in the Department of Chemical and Biomolecular Engineering, University of

Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

e-mail: discher@seas.upenn.edu

1. Shum, H. C., Zhao, Y., Kim, S.-H. & Weitz, D. A. *Angew. Chem. Int. Edn* **50**, 1648–1651 (2011).
2. Hayashi, T. & Carthew, R. W. *Nature* **431**, 647–652 (2004).
3. Manning, M. L. *et al. Proc. Natl Acad. Sci. USA* **107**, 12517–12522 (2010).
4. Discher, D. E. & Eisenberg, A. *Science* **297**, 967–973 (2002).
5. Christian, D. A. *et al. Nature Mater.* **8**, 843–849 (2009).
6. Kim, Y. *et al. J. Controlled Release* **134**, 132–140 (2009).

HOST–PATHOGEN INTERACTION

Culprit within a culprit

The parasitic infection mucocutaneous leishmaniasis can vary in severity. It emerges that the levels of an RNA virus within the parasite affect both the host's immune response and the parasite's persistence.

MARTIN OLIVIER

Protozoan parasites of the genus *Leishmania* are responsible for the infectious disease leishmaniasis. The disease mainly affects those living in tropical and sub-tropical parts of the world, and a unique form of it called mucocutaneous leishmaniasis causes a disfiguring disease in the Americas. Despite the century-old knowledge that *Leishmania* parasites of the *Viannia* subgenus cause leishmaniasis, how the disease progresses remains enigmatic. How does a primary cutaneous lesion lead to a secondary, metastasizing form reappearing months later in the nasopharyngeal tissues? Writing in *Science*, Ives *et al.*¹ shed light on this process, showing that the parasite recruits help from within to counteract the host's immune responses, thus enabling it to spread farther afield.

Mucocutaneous leishmaniasis is associated with a persistent inflammatory response characterized by increased expression of pro-inflammatory mediators (TNF- α and some chemokines)^{2,3} that play a pivotal part in recruiting immune-system cells such as macrophages to the site of infection. Using previous approaches^{4,5}, Ives *et al.* isolated clones of *Leishmania guyanensis* — a parasite of the *Viannia* subgenus — from infected hamsters.

The clones were either non-metastatic (*L.g.M*–) or metastatic (*L.g.M*+), and Ives *et al.* found that the *L.g.M*+ clones harboured a virus called *Leishmania* RNA virus 1 (LRV1). When the authors infected mouse macrophages with these clones, the cells rapidly expressed several cytokines and chemokines (including TNF- α and CXCL10) that are relevant to the pathology of mucocutaneous leishmaniasis. The strongest response was to *L.g.M*+ clones and to another clone isolated from patients

with mucocutaneous leishmaniasis. By contrast, *L.g.M*– clones and their corresponding human isolates from cutaneous lesions led to only a modest, albeit significant, response compared with uninfected cells; this is consistent with earlier findings comparing different *Leishmania* species⁶.

Ives *et al.*¹ show that the ability of the *L.g.M*+ clones to induce increased expression of pro-inflammatory molecules depends on their internalization into macrophages and their sequestration within phagolysosomes. A phagolysosome is an organelle in which engulfed pathogens can be destroyed; *Leishmania*, however, has evolved to survive its harsh environment.

An inflammatory immune response is often triggered by the recognition of a pathogen by specific receptors of the innate immune system. One such receptor family comprises the Toll-like receptors (TLRs), which are found on the cell surface or on vesicle membranes of macrophages and other cells of innate immunity. To determine how the inflammatory response to *L. guyanensis* comes about, Ives and colleagues studied macrophages that were functionally deficient in the various phagosomal TLRs (TLR3, 7 and 9) or in intracellular adaptors related to TLR signalling (MyD88 and TRIF).

They found that TLR3 and TRIF are essential for maximal production not just of pro-inflammatory mediators induced by *L.g.M*+, but also of IFN- β — a cytokine produced in response to TLR3 activation and which potentially can cause organ damage. Intriguingly, previous work on virus-free *Leishmania* species suggested that IFN- β ⁷ and TLR3⁸ could be involved, respectively, in the progression of leishmaniasis and in the parasite's recognition by IFN- γ -primed macrophages. Those studies

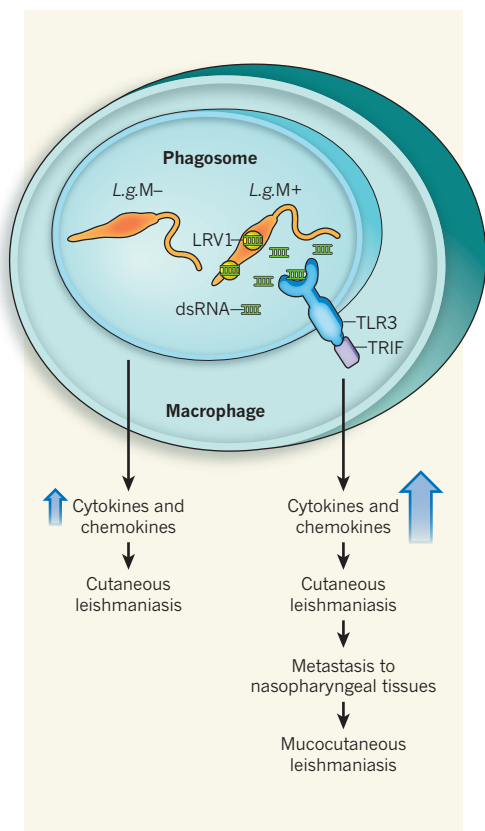


Figure 1 | What a difference RNA makes. Ives *et al.*¹ find that virus-infected *Leishmania* parasites (*L.g.M+*) trigger hyperinflammation through release of a viral double-stranded RNA (dsRNA) that binds to the complex of phagosomal receptor TLR3 and its adapter protein TRIF within macrophages. *Leishmania guyanensis* (*L.g.M-*) usually causes cutaneous ulceration that heals as a result of an effective immune response. However, the *L.g.M+* parasites infected with *Leishmania* RNA virus 1 (LRV1) can persist and metastasize to nasopharyngeal tissues. There they cause debilitating disfigurement due to an intense inflammatory response. Blue arrows indicate the levels of cytokines and chemokines.

did not, however, address the question of how the parasite modulates the host's inflammatory response.

Ives *et al.* find that, in metastasizing *L.g.M+*, LRV1 is responsible for the TLR3-dependent macrophage activation. TLR3, TLR7 and TLR9 recognize pathogenic nucleic-acid motifs, including double-stranded RNA (dsRNA). Indeed, purified dsRNA from LRV1 was as powerful as metastasizing parasites in inducing secretion of pro-inflammatory mediators, and this effect was completely abrogated in TLR3-deficient macrophages (Fig. 1).

This result is the cornerstone of the authors' study, revealing for the first time how a virus within a pathogen can modify its pathological development within the host. Nonetheless, although the authors report that TLR3-deficient mice have a lower parasite burden and make a reduced inflammatory response (as

evidenced by a decreased response in the footpad-swelling test compared with wild-type animals), they found no significant difference in the development of mucocutaneous leishmaniasis.

A disadvantage of studying mucocutaneous leishmaniasis in mice is that these animals do not develop the nasopharyngeal pathology. It would be better to investigate the effects of TLR3 deficiency in hamsters to determine how *L.g.M+* migrates from a primary lesion to the nasopharyngeal tissues to cause inflammation-mediated tissue destruction.

Ives *et al.*¹ hypothesize that, early in the course of an infection, LRV1 dsRNA is released from dead parasites that could not survive within macrophages. Could this be so? The same team reported previously⁹ that *L.g.M+* parasites are more resistant to reactive oxygen species and nitric oxide, which would make them better at resisting the harsh environment inside macrophage phagosomes compared with parasites that are not infected with viruses. This could be why, despite *L.g.M+* parasites inducing a stronger immune response, *L.g.M+*-infected animals develop greater footpad swelling and a higher parasite burden.

Another question is whether the aflagellate intracellular form of the *Leishmania* parasite, which can migrate to the host's nasopharyngeal

tissues, is as effective at inducing an inflammatory response as its predecessor flagellate form, which is inoculated into the host by the sand-fly vector. Without knowledge of the cellular and molecular mechanisms underlying the tropism of *L.g.M+* parasites, how mucocutaneous leishmaniasis progresses remains a mystery. Nevertheless, Ives and co-workers' paper¹ is important for providing the first evidence of how a 'pathogen' within a pathogen can modify the course of an infection. ■

Martin Olivier is at the Research Institute of the McGill University Health Centre, Department of Microbiology and Immunology, Montreal, Quebec H3G 1A4, Canada. e-mail: martin.olivier@mcgill.ca

1. Ives, A. *et al.* *Science* **331**, 775–778 (2011).
2. Pirmez, C. *et al.* *J. Clin. Invest.* **91**, 1390–1395 (1993).
3. Vargas-Inchaustegui, D. A. *et al.* *Infect. Immun.* **78**, 301–308 (2010).
4. Travi, B. *et al.* *J. Parasitol.* **74**, 1059–1062 (1988).
5. Martinez, J. E. *et al.* *J. Parasitol.* **86**, 792–799 (2000).
6. Forget, G. *et al.* *Eur. J. Immunol.* **35**, 1906–1917 (2005).
7. Mattner, J. *et al.* *J. Immunol.* **172**, 7574–7582 (2004).
8. Flandin, J. F. *et al.* *Eur. J. Immunol.* **36**, 411–420 (2006).
9. Acestor, N. *et al.* *J. Infect. Dis.* **194**, 1160–1167 (2006).

EARTHQUAKES

A Chilean surprise

Initiation of the great 2010 Chile earthquake occurred within the rupture zone of the 1835 event experienced by Charles Darwin. However, the peak fault slip was to the north of the epicentre — not where it was expected to occur.

THORNE LAY

On 20 February 1835, Charles Darwin was resting in the woods near Valdivia, Chile, when the ground suddenly trembled for about two minutes, although “the time appeared much longer”. Two weeks later, at Concepción, Darwin saw the devastation wrought by the ground shaking and tsunami, and observed coastal uplift that he inferred to have “accompanied or caused the earthquake”. This great earthquake, with an estimated² seismic magnitude (M_w) of about 8.5, is now understood to have occurred on a megathrust fault, where the Nazca plate underthrusts the South American plate with a plate-convergence rate of about 70 millimetres per year.

After 175 years, the region ruptured again, in the great M_w -8.8 earthquake of 27 February 2010, again causing extensive shaking and tsunami damage. Studies by Moreno *et al.*³ and

Lorito *et al.*⁴ have respectively quantified the fault-zone processes before and after the 2010 earthquake, revealing surprising contrasts to the event witnessed by Darwin.

Repeated stick-slip sliding of plate-boundary faults occurs as a result of the frictional build-up and release of stress driven by relative plate motions. In 1939, Concepción was again struck by a devastating earthquake, but the faulting occurred within the sinking Nazca plate, not on the plate boundary². The plate boundary north of the 1835 rupture experienced large underthrusting earthquakes in 1906, 1928 and 1985, whereas to the south the immense M_w -9.5 earthquake of 1960 accommodated plate convergence in that region (Fig. 1a). Between 1835 and 2010, some 12 metres of relative plate convergence had occurred near the Darwin earthquake zone, presumably with frictional resistance causing extensive build-up of elastic strain in the rock around the fault zone that was released in the

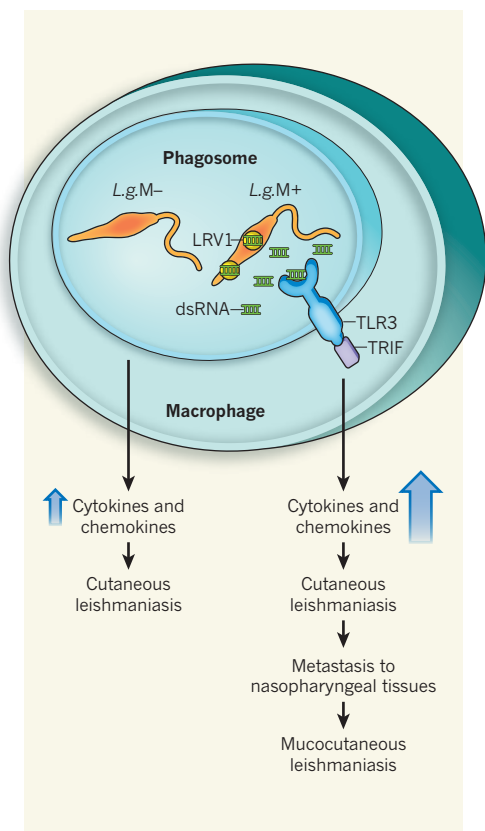


Figure 1 | What a difference RNA makes. Ives *et al.*¹ find that virus-infected *Leishmania* parasites (*L.g.M+*) trigger hyperinflammation through release of a viral double-stranded RNA (dsRNA) that binds to the complex of phagosomal receptor TLR3 and its adapter protein TRIF within macrophages. *Leishmania guyanensis* (*L.g.M-*) usually causes cutaneous ulceration that heals as a result of an effective immune response. However, the *L.g.M+* parasites infected with *Leishmania* RNA virus 1 (LRV1) can persist and metastasize to nasopharyngeal tissues. There they cause debilitating disfigurement due to an intense inflammatory response. Blue arrows indicate the levels of cytokines and chemokines.

did not, however, address the question of how the parasite modulates the host's inflammatory response.

Ives *et al.* find that, in metastasizing *L.g.M+*, LRV1 is responsible for the TLR3-dependent macrophage activation. TLR3, TLR7 and TLR9 recognize pathogenic nucleic-acid motifs, including double-stranded RNA (dsRNA). Indeed, purified dsRNA from LRV1 was as powerful as metastasizing parasites in inducing secretion of pro-inflammatory mediators, and this effect was completely abrogated in TLR3-deficient macrophages (Fig. 1).

This result is the cornerstone of the authors' study, revealing for the first time how a virus within a pathogen can modify its pathological development within the host. Nonetheless, although the authors report that TLR3-deficient mice have a lower parasite burden and make a reduced inflammatory response (as

evidenced by a decreased response in the footpad-swelling test compared with wild-type animals), they found no significant difference in the development of mucocutaneous leishmaniasis.

A disadvantage of studying mucocutaneous leishmaniasis in mice is that these animals do not develop the nasopharyngeal pathology. It would be better to investigate the effects of TLR3 deficiency in hamsters to determine how *L.g.M+* migrates from a primary lesion to the nasopharyngeal tissues to cause inflammation-mediated tissue destruction.

Ives *et al.*¹ hypothesize that, early in the course of an infection, LRV1 dsRNA is released from dead parasites that could not survive within macrophages. Could this be so? The same team reported previously⁹ that *L.g.M+* parasites are more resistant to reactive oxygen species and nitric oxide, which would make them better at resisting the harsh environment inside macrophage phagosomes compared with parasites that are not infected with viruses. This could be why, despite *L.g.M+* parasites inducing a stronger immune response, *L.g.M+*-infected animals develop greater footpad swelling and a higher parasite burden.

Another question is whether the aflagellate intracellular form of the *Leishmania* parasite, which can migrate to the host's nasopharyngeal

tissues, is as effective at inducing an inflammatory response as its predecessor flagellate form, which is inoculated into the host by the sand-fly vector. Without knowledge of the cellular and molecular mechanisms underlying the tropism of *L.g.M+* parasites, how mucocutaneous leishmaniasis progresses remains a mystery. Nevertheless, Ives and co-workers' paper¹ is important for providing the first evidence of how a 'pathogen' within a pathogen can modify the course of an infection. ■

Martin Olivier is at the Research Institute of the McGill University Health Centre, Department of Microbiology and Immunology, Montreal, Quebec H3G 1A4, Canada. e-mail: martin.olivier@mcgill.ca

1. Ives, A. *et al.* *Science* **331**, 775–778 (2011).
2. Pirmez, C. *et al.* *J. Clin. Invest.* **91**, 1390–1395 (1993).
3. Vargas-Inchaustegui, D. A. *et al.* *Infect. Immun.* **78**, 301–308 (2010).
4. Travi, B. *et al.* *J. Parasitol.* **74**, 1059–1062 (1988).
5. Martinez, J. E. *et al.* *J. Parasitol.* **86**, 792–799 (2000).
6. Forget, G. *et al.* *Eur. J. Immunol.* **35**, 1906–1917 (2005).
7. Mattner, J. *et al.* *J. Immunol.* **172**, 7574–7582 (2004).
8. Flandin, J. F. *et al.* *Eur. J. Immunol.* **36**, 411–420 (2006).
9. Acestor, N. *et al.* *J. Infect. Dis.* **194**, 1160–1167 (2006).

EARTHQUAKES

A Chilean surprise

Initiation of the great 2010 Chile earthquake occurred within the rupture zone of the 1835 event experienced by Charles Darwin. However, the peak fault slip was to the north of the epicentre — not where it was expected to occur.

THORNE LAY

On 20 February 1835, Charles Darwin was resting in the woods near Valdivia, Chile, when the ground suddenly trembled for about two minutes, although “the time appeared much longer”. Two weeks later, at Concepción, Darwin saw the devastation wrought by the ground shaking and tsunami, and observed coastal uplift that he inferred to have “accompanied or caused the earthquake”. This great earthquake, with an estimated² seismic magnitude (M_w) of about 8.5, is now understood to have occurred on a megathrust fault, where the Nazca plate underthrusts the South American plate with a plate-convergence rate of about 70 millimetres per year.

After 175 years, the region ruptured again, in the great M_w -8.8 earthquake of 27 February 2010, again causing extensive shaking and tsunami damage. Studies by Moreno *et al.*³ and

Lorito *et al.*⁴ have respectively quantified the fault-zone processes before and after the 2010 earthquake, revealing surprising contrasts to the event witnessed by Darwin.

Repeated stick-slip sliding of plate-boundary faults occurs as a result of the frictional build-up and release of stress driven by relative plate motions. In 1939, Concepción was again struck by a devastating earthquake, but the faulting occurred within the sinking Nazca plate, not on the plate boundary². The plate boundary north of the 1835 rupture experienced large underthrusting earthquakes in 1906, 1928 and 1985, whereas to the south the immense M_w -9.5 earthquake of 1960 accommodated plate convergence in that region (Fig. 1a). Between 1835 and 2010, some 12 metres of relative plate convergence had occurred near the Darwin earthquake zone, presumably with frictional resistance causing extensive build-up of elastic strain in the rock around the fault zone that was released in the

great 2010 Chile earthquake. Such a build-up of strain can now be detected by modern geodetic methods.

The work of Moreno *et al.*³ consisted of analysis of a decade of ground-deformation data prior to 2010, recorded by a dense network of Global Positioning System (GPS) sensors along the length of Chile. These data indicated that the western margin of the South American plate north of the 1960 rupture zone had generally eastward ground velocities relative to the stable interior of the plate — a manifestation of strong frictional coupling between the converging Nazca and South American plates, and of accumulating strain in the overriding plate.

The authors' finite-element modelling³ of the geodetic velocities as the result of interseismic 'shortening' of the Andean subduction zone yielded estimates of the spatially varying interplate coupling of the plate-boundary megathrust fault (Fig. 1b). They found that the southern half (35.5° S–37.5° S) of the 1835 rupture zone was 'locked' (not slipping). The northern region of the rupture zone seemed to be 'partially locked' between 34.8° S and 35.5° S, meaning that some interplate sliding was occurring by non-seismic slow slip, although not keeping up with the total plate convergence rate, effectively lessening the accumulated strain and slip deficit by up to 50%. The megathrust was not slipping farther north, from 33.8° S to 34.8° S, along the 1985 rupture zone and the northern portion of the 1928 rupture zone.

With the megathrust regions both north and south of the 1835 rupture having experienced relatively recent large ruptures, the accumulated stress on those areas would presumably be lower than in the strongly locked portion of the 1835 zone. On the basis of plate-tectonic convergence rate, the history of large underthrusting earthquakes, and geodetically inferred interplate locking for the previous decade, one might reasonably expect the slip distribution of the 2010 event to have been confined to the region between the 1960 and 1928 ruptures, with maximum slip in the southern portion. However, determinations of the actual slip distribution, including that by Lorito *et al.*⁴, are largely incompatible with these expectations.

The great 2010 Chile earthquake did initiate in the central deeper portion of the 1835 rupture zone. But the rupture expanded bilaterally over a total length of about 500 km, with the aftershock zone overlapping the 1928 and 1985 rupture zones and the northernmost part of the 1960 rupture zone (Fig. 1b). The event was thus much larger than the 1835 event, possibly resembling an earlier earthquake in 1751. Persistent segmentation of megathrust ruptures along the coast is clearly not pronounced in this subduction zone.

Comparison of initial rupture models for the megathrust slip distribution, made on the basis of remotely recorded seismic waves,

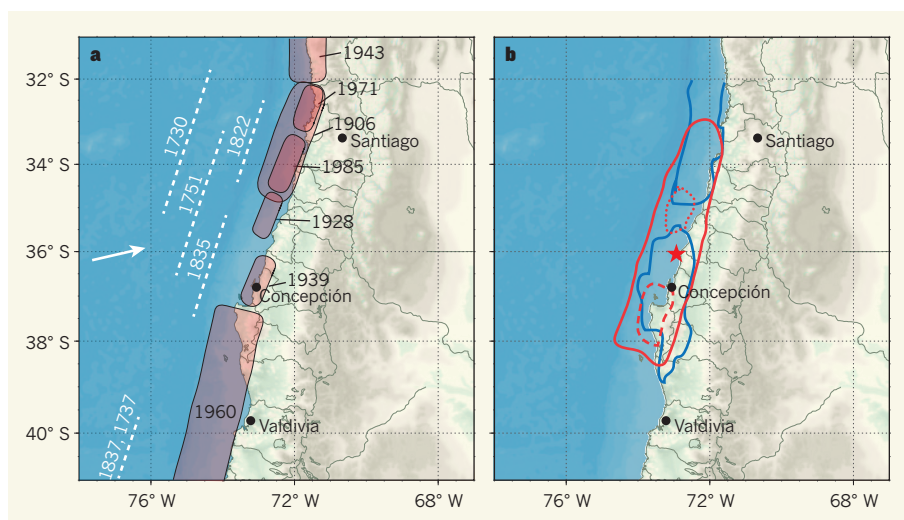


Figure 1 | Earthquake history and the rupture zone of the great 27 February 2010 Chile earthquake. **a**, Source regions of large earthquakes over the past century along the central Chile subduction zone, shown as outlined areas with corresponding event dates. This subduction zone is where the Nazca plate (to the left) thrusts under the South American plate (to the right). The white arrow indicates the plate-convergence direction. The approximate rupture lengths for events before 1900 are indicated by white dashed lines, with the 1835 event being the Darwin earthquake. **b**, Inferences of fault-zone processes for before³ and after⁴ the 2010 event. The blue curves indicate regions of >75% locking of the megathrust, as estimated³ from a decade of GPS observations prior to 2010. The red curves depict post-event analysis of the earthquake's behaviour⁴. Solid red, the 2010 earthquake aftershock zone; dotted red, the region with unexpected slip greater than 15 m north of the 2010 epicentre (red star); dashed red, a region with slip of 5–10 m south of the epicentre.

showed a strong correlation with the geodetically inferred regions of strong locking^{3,5}. But refined models — which involved matching data from GPS displacements, tsunami recordings and radar-remote-sensing observations, and measurements of coastal uplift^{4,6,7} — reduce the correlation by repositioning the main slip patches by 30–50 km.

These models place the largest slip, of up to 20 m, north of the epicentre at 34.5° S–35.5° S, overlapping the region of previously partial (50–75%) locking in the 1928 rupture zone, where the pre-stress was presumably relatively low. Co-seismic slip of 5–10 m did occur in the southern region of the 1835 rupture zone that previously had strong locking, but the strongly locked region of the megathrust west of the epicentre has only a few metres of slip in all of the refined models^{4,6,7}. Lorito *et al.*⁴ thus infer that there have been large increases in shear stress on the locked megathrust in this region, and so there is continuing potential for a near-future large earthquake in the 1835 rupture zone. However, our lack of understanding of why the region did not rupture during the 2010 earthquake casts doubt on our ability to project its future behaviour.

The surprising size and slip distribution of the 2010 Chile earthquake has now come into focus, highlighting the shortcomings in our ability to anticipate the rupture properties of great earthquakes, even along plate boundaries where we can quantify interseismic locking. Like the 26 December 2004 (M_w 9.2) Sumatra earthquake⁸, the 2010 rupture extended

far beyond conventional expectations. The northward expansion of rupture beyond the 1835 rupture zone may reflect the intrinsic tendency of ruptures to continue as long as no acute barrier or strain deficit causes them to arrest.

The patchiness of the co-seismic slip in the 2010 rupture emphasizes the importance of quantifying large-rupture slip distributions, because they will influence the pre-stress conditions for future earthquakes. It seems that the region of Darwin's earthquake still has the potential to make “the earth tremble”, but when it will next do so remains uncertain. ■

Thorne Lay is in the Department of Earth and Planetary Sciences, University of California Santa Cruz, Santa Cruz, California 95064, USA.

e-mail: tlay@ucsc.edu

1. Darwin, C. *Journal of Researches into the Natural History and Geology of the Countries Visited During the Voyage of H.M.S. Beagle Round the World* 2nd edn (Murray, 1845).
2. Beck, S., Barrientos, S., Kausel, E. & Reyes, M. *J. South Am. Earth Sci.* **11**, 115–129 (1998).
3. Moreno, M., Rosenau, M. & Oncken, O. *Nature* **467**, 198–202 (2010).
4. Lorito, S. *et al.* *Nature Geosci.* **4**, 173–177 (2011).
5. Lay, T. *et al.* *Geophys. Res. Lett.* **37**, L13301, doi:10.1029/2010GL043379 (2010).
6. Tong, X. *et al.* *Geophys. Res. Lett.* **37**, L24311, doi:10.1029/2010GL045805 (2010).
7. Delouis, B., Nocquet, J.-M. & Vallée, M. *Geophys. Res. Lett.* **37**, L17305, doi:10.1029/2010GL043899 (2010).
8. Lay, T. *et al.* *Science* **308**, 1127–1133 (2005).

Network anatomy and *in vivo* physiology of visual cortical neurons

Davi D. Bock^{1,2}, Wei-Chung Allen Lee^{1,2}, Aaron M. Kerlin¹, Mark L. Andermann¹, Greg Hood³, Arthur W. Wetzel³, Sergey Yurgenson¹, Edward R. Soucy², Hyon Suk Kim^{1,2} & R. Clay Reid^{1,2}

In the cerebral cortex, local circuits consist of tens of thousands of neurons, each of which makes thousands of synaptic connections. Perhaps the biggest impediment to understanding these networks is that we have no wiring diagrams of their interconnections. Even if we had a partial or complete wiring diagram, however, understanding the network would also require information about each neuron's function. Here we show that the relationship between structure and function can be studied in the cortex with a combination of *in vivo* physiology and network anatomy. We used two-photon calcium imaging to characterize a functional property—the preferred stimulus orientation—of a group of neurons in the mouse primary visual cortex. Large-scale electron microscopy of serial thin sections was then used to trace a portion of these neurons' local network. Consistent with a prediction from recent physiological experiments, inhibitory interneurons received convergent anatomical input from nearby excitatory neurons with a broad range of preferred orientations, although weak biases could not be rejected.

Past studies of the synaptic connections in the cortex have focused on a few cells at a time, so that today we have only partial information about the structure of highly interconnected cortical networks. Until recently, anatomists relied on the sparse labelling of individual neurons to trace the extensive axonal and dendritic arbors of cortical neurons. Synaptic connectivity was originally inferred from the overlap of axonal and dendritic arbors¹, an approach that remains fruitful today^{2,3}. Combined physiological and anatomical studies of cortical slices *in vitro* have extended this analysis to include rules of pair-wise connectivity between different cell types and cortical layers^{4–6}. It has been argued that the cortical network might be random beyond these simple statistical rules⁷, but recent studies suggest that there could be higher-order patterns of connections in cortical networks, such as mutually interconnected triplets of cells⁸, and subnetworks of neurons that are highly connected within a group, but not between groups^{9,10}. Although these hypothesized patterns are based on *in vitro* physiological data, they have never been demonstrated anatomically, nor have they been related to information processing as measured from neural response properties *in vivo*.

Electron microscopy (EM) is an ideal tool for characterizing the highly interconnected structure of cortical networks. From visualizing quantal release by synaptic vesicles¹¹ to generating a complete wiring diagram of a model organism¹², it has provided definitive data for examining the relationship between structure and function in the nervous system. Electrophysiology combined with light microscopy and serial-section EM has allowed the inspection of structure–function relationships of single cells within neural circuits, such as the hippocampus¹³, retina^{14,15}, thalamus¹⁶ and cortex^{17–20}. A crucial difference from light microscopy is that serial-section EM can be used to follow the three-dimensional contours of neuronal membranes, so that any given axon or dendrite can be traced over hundreds of micrometres without selective, sparse staining of individual neurons. This feature of serial-section EM has been used to examine small volumes of cortical tissue, typically numbering in the thousands of cubic micrometres, in which portions of multiple dendrites and axons were reconstructed to

examine synaptic relationships amongst them (refs 21, 22, compare ref. 23).

Here we exploited recent improvements in computer speed and storage capacity to perform serial-section EM of a volume that encompasses millions of cubic micrometres, sufficient to contain large portions of the dendritic and axonal arbors of more than 1,000 cells. With this data set, we could attempt a sampling—targeted to a subset of functionally imaged cells—of the dense interconnections found in a cortical network. In particular, we tested a somewhat controversial prediction from recent physiological work (refs 24–27, compare refs 28, 29): that inhibitory interneurons in the mouse primary visual cortex receive dense, convergent input from nearby excitatory (pyramidal) neurons with widely varying preferred stimulus orientations (Fig. 1a).

Functional imaging and large-scale EM

We used *in vivo* two-photon calcium imaging to determine the preferences for stimulus orientation of a cohort of cells in layer 2/3 of mouse primary visual cortex^{27,30}. After loading cells with the fluorescent calcium indicator Oregon Green BAPTA-1 AM (OGB)³¹, we recorded neuronal activity as reflected by increases in intracellular calcium from a single plane of cells 186 μm beneath the brain surface. Black and white bars of varying orientations and directions were presented to the anaesthetized animal, and the cellular responses were used to generate a map of orientation preference (Fig. 1b and Supplementary Fig. 1; Methods). We concluded the *in vivo* experiment by collecting a stack of images from a volume surrounding the calcium-imaged plane (Fig. 1c; red, vessels; green, neurons). We then immediately perfused the animal and prepared the corresponding volume of visual cortex for serial-section transmission EM (TEM; Methods).

After finding the calcium-imaged region of cortex (Methods), we cut a series of 1,215 thin sections (40–45 nm), oriented radially, at right angles to the functionally imaged plane (Fig. 1b, c). Each section was wide enough (450 μm) to encompass the imaged plane, and tall enough (350 μm) to include cortical layers 1, 2/3 and upper 4 (Fig. 2a, b, Supplementary Movie 1). Although cortical axons can

¹Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA. ²The Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA. ³National Resource for Biomedical Supercomputing, Pittsburgh Supercomputing Center, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA.

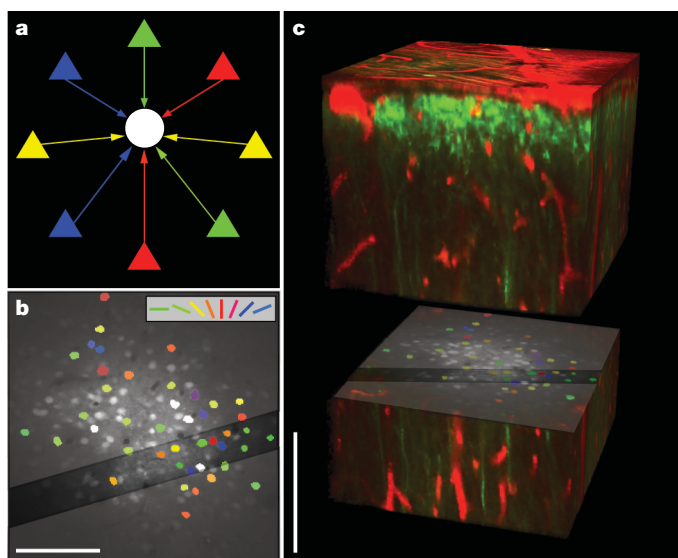


Figure 1 | Functional characterization of neurons before anatomical reconstruction. **a**, Schematic representation of diverse input to inhibitory interneurons. Excitatory pyramidal cells (triangles) with varied preferred orientations (different colours) provide input (coloured arrows) to an inhibitory interneuron (white circle). **b**, Cell-based visual orientation preference map in the mouse visual cortex from *in vivo* two-photon calcium imaging. Visually responsive cells are coloured according to their estimated preferred orientation (colour coding shown at top), with broadly tuned cells (orientation selectivity index ≤ 0.2) shown white. Dark diagonal band, region targeted for acquisition of EM sections, cut orthogonal to image plane. **c**, *In vivo* two-photon fluorescence image of the three-dimensional volume (red, blood vessels or astrocytes; green, OGB-loaded somata or yellow fluorescent protein (YFP)-labelled apical dendrites) separated to expose the functionally imaged plane. Scale bars, 100 μm .

travel for millimetres, the proximal axonal collaterals of pyramidal cells arborize more locally⁷, such that a portion of their local connectivity could be sampled within the span of the EM volume. In line with previous work, we found that in order to reliably trace the finest neural processes and to identify chemical synapses (Fig. 2f, Supplementary Movie 2), we needed <50 nm section thickness and <5 nm lateral resolution³².

We built a custom TEM Camera Array (TEMCA) that used four high-speed CCD (charge-coupled device) cameras to efficiently acquire EM images of the required size and resolution (Supplementary Fig. 2; Methods). The TEMCA achieves an order of magnitude increase in throughput over most commercially available TEM imaging systems. This system enabled us to collect the thousands of individual images required to mosaic each of the serial sections in approximately 20 min (Fig. 2a), so that we could image the entire series over the course of several months.

The 3.2×10^6 camera images in our data set were converted into a seamless three-dimensional image volume by globally registering and aligning the EM volume after automated stitching of adjacent images into sections (Methods). The raw image data were approximately 36×10^{12} bytes (36 terabytes, TB), with the final stitched and registered EM data set encompassing approximately ten million megapixels (10 TB), spanning $\sim 450 \times 350 \times 52 \mu\text{m}$ and containing approximately 1,500 cell bodies (Fig. 2b).

We next re-located in the EM volume the cells whose function had been characterized by *in vivo* calcium imaging. The large extent of the EM data set permitted registration of the fluorescently labelled vasculature and neurons between two-photon and EM imaging (Fig. 3a–c). We matched blood vessels of successively finer calibre, followed by the cell bodies, in order to identify individual neurons of known orientation selectivity within the EM volume (Fig. 3d, Supplementary Fig. 3; Methods). The EM volume intersected the cell bodies of 14 visually

responsive neurons (Fig. 1 and Supplementary Fig. 1). Thirteen of the cells were selective for stimulus orientation. The fourteenth neuron was visually responsive, but non-selective for orientation. It had a non-pyramidal neuronal morphology, received asymmetric contacts onto the cell body, and made symmetric synapses onto its postsynaptic targets (Supplementary Fig. 4), suggesting it was an inhibitory, GABAergic interneuron³³. Both the precision of physical registration and the alignment of appropriate functional properties with neuronal morphology (orientation selective excitatory cells and an unselective inhibitory neuron) demonstrate that we successfully combined micrometre-scale *in vivo* functional imaging and nanometre-scale EM ultrastructure.

EM circuit reconstruction and the resultant network graph

To analyse the circuit's anatomical connectivity, we manually traced a wire-frame model of the dendritic and axonal arbors of each functionally characterized neuron, and noted the location of each synapse along the axons. For each synapse, we then reconstructed the postsynaptic dendrite centripetally until we reached either the cell body or the boundary of the EM volume (Methods). In this directed manner, we were able to identify all of the postsynaptic targets that could be found in the EM-imaged volume and to determine when multiple axons converged onto a common postsynaptic target, without unnecessary tracing of the target cells' dendrites (Fig. 4, Supplementary Fig. 5).

We categorized the postsynaptic targets as either excitatory or inhibitory on the basis of morphology. Pyramidal cell dendrites were densely studded with spines, whereas inhibitory interneuron dendrites were sparsely spinous and receive more asymmetric (excitatory) synapses their dendritic shafts³³. Of 245 synapses originating from 10 functionally characterized pyramidal neurons that made synapses in the EM-imaged volume, 125 (51%) were onto inhibitory dendrites and 120 (49%) onto excitatory dendrites. Of the 185 distinct postsynaptic targets, 71 (38%) were inhibitory (Fig. 5, cyan) and 114 (62%) were excitatory (Fig. 5, magenta). The proportion of inhibitory cells in the target population was less than the proportion of synapses onto inhibitory targets because individual axons often made multiple synapses with inhibitory targets. We rendered the anatomical reconstructions of the functionally characterized cells and their postsynaptic targets as a graph (Fig. 5b) to examine the functional logic of the network.

Convergent excitatory input to inhibitory interneurons

In the graph of connectivity (Fig. 5b), we found multiple examples of pyramidal cells with diverse preferred stimulus orientations that provided convergent input to inhibitory neurons. We restricted subsequent analysis to a reduced and verified subnetwork that included convergent connections, as well as connections amongst the functionally imaged neurons (Fig. 5c, d, Supplementary Figs 6, 7; Methods). Most strikingly, some inhibitory targets received input from three or four distinct cells with a range of preferred orientations (Fig. 5d, Supplementary Fig. 7a–d).

When we examined all of the convergences onto inhibitory targets, most were from pairs of cells that were close to each other (Figs 5c, 6a), independent of the difference between their preferred orientations, which ranged from nearly orthogonal (Fig. 5d centre, Supplementary Fig. 7e–g) to nearly identical (Fig. 5d lower right, Supplementary Fig. 7m–o). The strongest predictor of whether two axons converged on a common target was found by examining how many of their synapses were nearby in space (Fig. 6b, c; cumulative synaptic proximity, $P < 1.3 \times 10^{-5}$; Methods).

The functional properties of a pair of excitatory cells, specifically the difference in their preferred orientations, was not predictive of whether they converged onto an inhibitory target. Pairs that converged had a distribution indistinguishable from the uniform distribution (Fig. 6d, red line; $P > 0.30$) and from a distribution in which all possible presynaptic pairs were considered, weighted by their proximity (Fig. 6d, blue line; $P > 0.68$; Methods). In summary, the axonal geometry of two excitatory neurons was a good predictor of whether

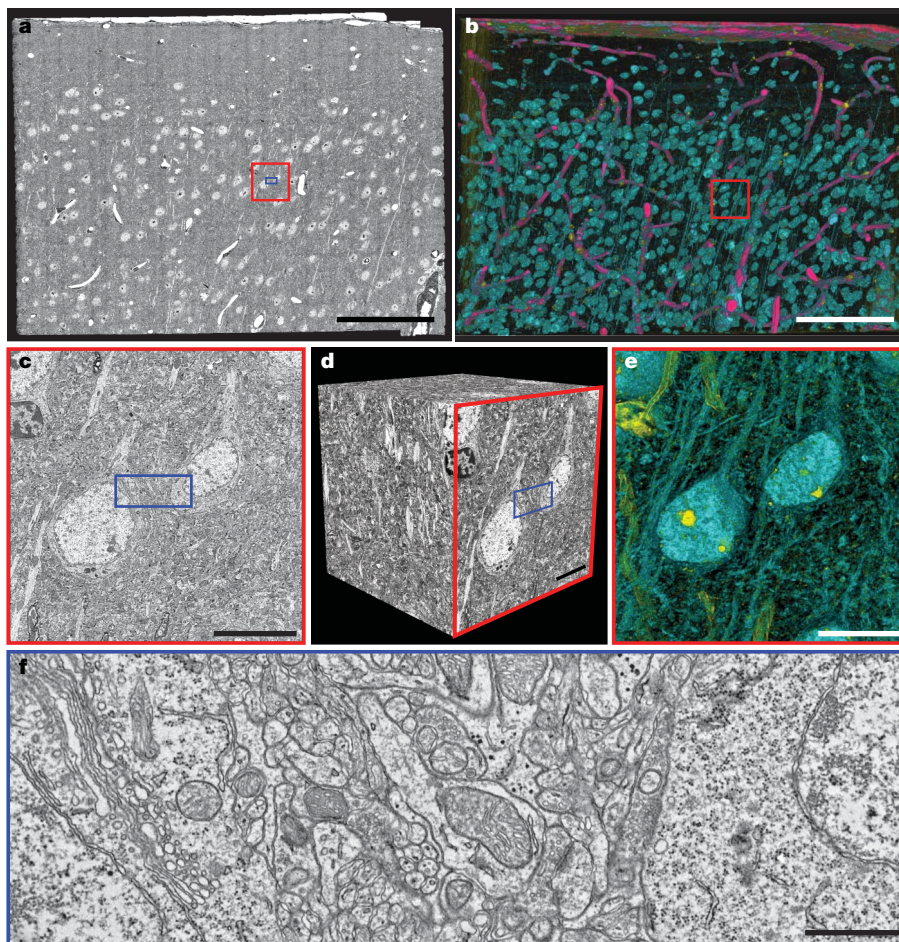


Figure 2 | Large-scale EM. **a**, Electron micrograph of an entire $120,000 \times 80,000$ pixel thin section, showing the pial surface (top) and cortical layers 1 through to upper 4 (bottom). **b**, **e**, Three-dimensional renderings of the EM volume through the entire series (**b**, Supplementary Movie 3) and through 50 sections of the cube in **d** (**e**, red outline in **b**, Supplementary Movie 4). **d**, A cube of the EM volume with **c** as one face. **c**, **f**, Zoomed-in view of two functionally characterized cells (**c**; red outline in **a**), and the neuropile between them (**f**; blue outlines in **a**, **c** and **d**), illustrating the density of axons and dendrites coursing between cell bodies. Pink represents electron transparent regions (for example, blood vessels), yellow represents regions that are electron dense (such as nucleoli, oligodendrocyte nuclei, and myelin), and aqua denotes regions with pixel values in between (for example, nuclei, cell bodies and dendrites). Scale bars: **a**, **b**, 100 μm ; **c**–**e**, 10 μm ; **f**, 1 μm .

they converged upon a nearby inhibitory interneuron, whereas the visual physiology of the excitatory neurons was not.

Discussion

The ability to study populations of neurons with a combination of network anatomy and *in vivo* physiology creates new opportunities for examining how neuronal circuits process information. Here we explored how both the geometry and the function of cortical neurons influence the patterns of connections between them. In the case of excitatory input to local inhibitory interneurons, geometry appeared to dominate over function⁷. This finding may provide an anatomical substrate for a prediction from recent physiological studies of mouse visual cortex, in which inhibitory neurons were found to be less selective than excitatory neurons (refs 24–27, compare refs 28, 29; Fig. 1a). Inhibitory interneurons that pool excitatory input could be

used to set the gain of orientation-selective pyramidal cells^{26,34,35}; they might also be involved in modulation of brain state³⁴ or in attention-dependent normalization of cortical activity³⁶.

Several studies have found that some subtypes of inhibitory neurons demonstrate orientation selectivity in the mouse^{28,29}. We did not classify different subtypes of inhibitory neurons, and thus cannot rule out the possibility that some might receive more selective input. Furthermore, although our sample size was large enough to exclude a strong bias in the preferred orientation tuning of convergent cell pairs, a larger sampling is necessary to exclude weaker biases (Methods).

Until it is possible to fully reconstruct large EM volumes^{22,37,38}, analysis of network connectivity will be limited to a partial sampling of the underlying anatomy. Here we concentrated on reconstructing the axons of functionally characterized pyramidal cells and their post-synaptic targets. Within our sample, we found that 51% of synapses

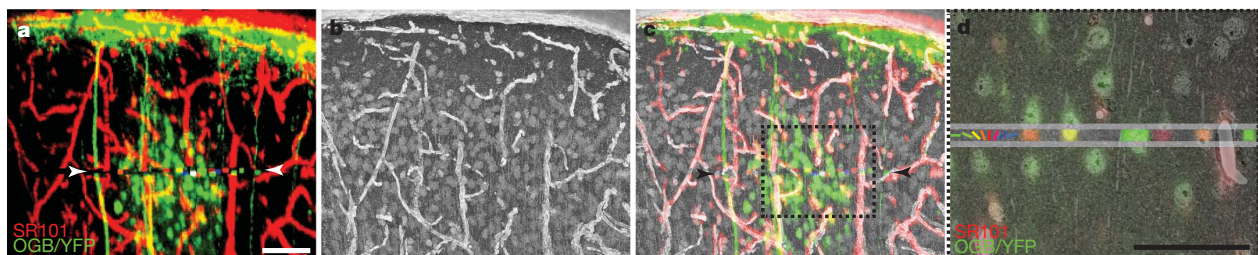


Figure 3 | Correspondence between *in vivo* fluorescence anatomy and EM. **a**, Maximum-intensity projection of the *in vivo* fluorescence anatomy corresponding to the EM volume (red, blood vessels or astrocytes; green, OGB or YFP as in Fig. 1c). Colours between arrowheads correspond to orientation preference of neurons in the imaged plane, as in **d**. **b**, Projection through 37 EM sections evenly spaced in the series. **c**, Merge of EM and fluorescence projections. **d**, Zoomed-in view of the region outlined in **c**, showing the overlay

of *in vivo* and EM data in a single thin section. Horizontal grey lines delineate the functionally imaged plane. Within the functionally imaged plane, the cell bodies of six neurons of known orientation preference (from left to right: cells 13, 12, 11, 10, 9 and 8, coloured as in Fig. 1b) are well registered with their EM ultrastructure. Outside the functionally imaged plane, blood vessels and astrocytes in the EM are well registered with the red SR101 staining. Scale bars, 50 μm .

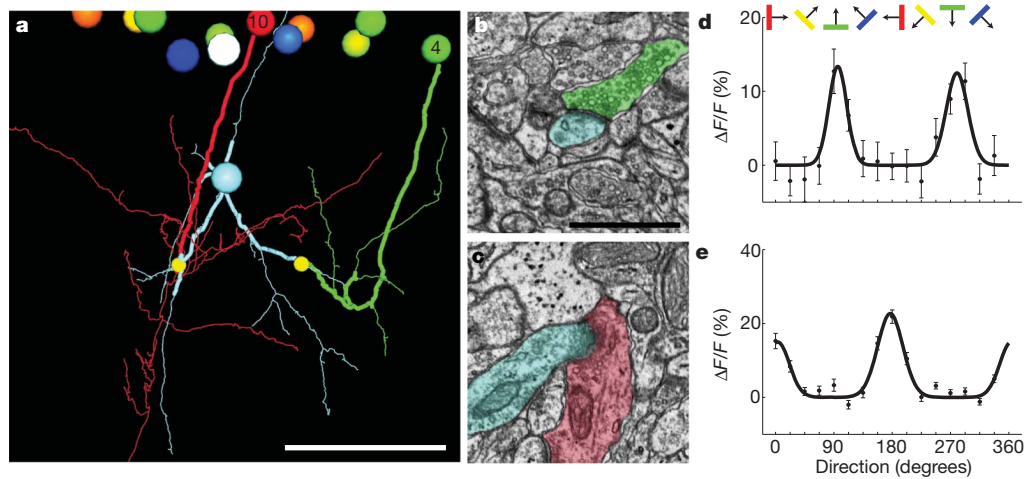


Figure 4 | Convergent synaptic input onto inhibitory interneurons.

a, Three-dimensional rendering of axonal contacts onto a postsynaptic neuron. Large balls at the top represent cell bodies of neurons within the functionally imaged plane. Axons of a horizontally tuned neuron (cell 4; green) and a vertically tuned neuron (cell 10; red) descend and make synapses (small yellow balls) onto dendrites of an inhibitory interneuron (cyan). The axonal and dendritic segments leading to the convergence were independently traced by a

second person, blind to the original segmentation (thick tracing). Cell bodies and axons coloured by orientation preference, as in Fig. 1b. Scale bar, 50 μ m. **b, c**, Electron micrographs showing the synapses onto the inhibitory neuron from cell 4 (**b**) and cell 10 (**c**) with corresponding colours overlaid. Scale bar, 1 μ m. **d, e**, Orientation tuning curves derived from *in vivo* calcium imaging of the cell bodies of cell 4 (**d**) and cell 10 (**e**). Coloured bars and arrows, stimulus orientation and direction. $\Delta F/F$, change in fluorescence. Error bars, \pm s.e.m.

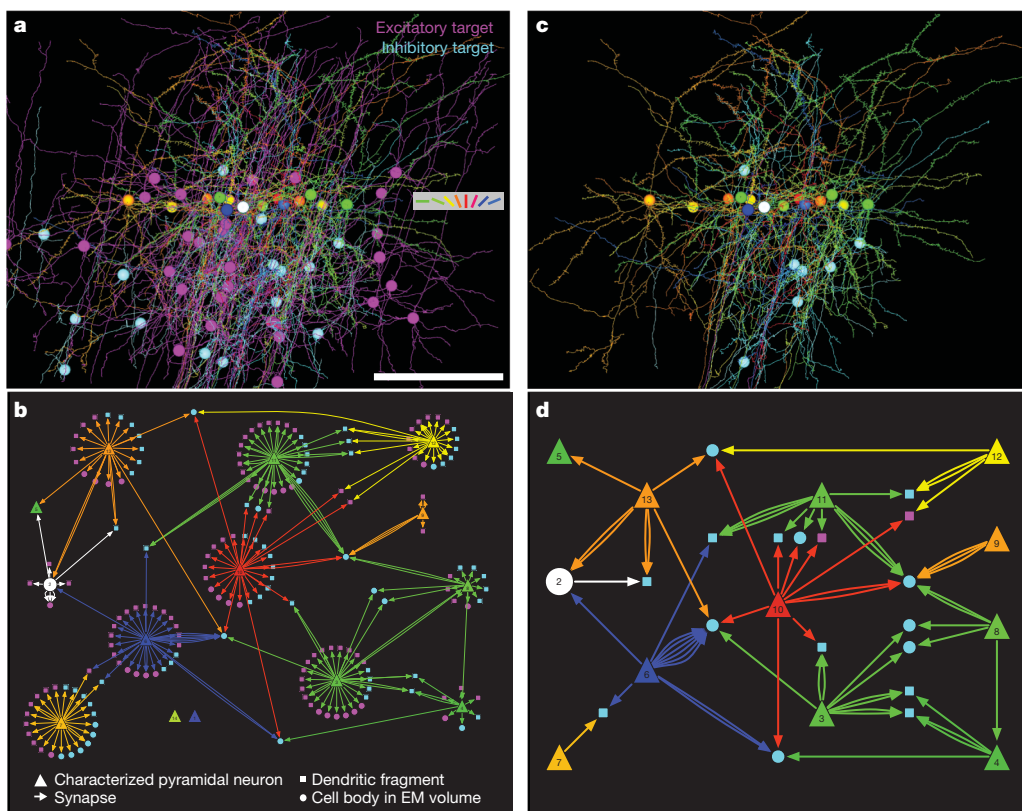


Figure 5 | From anatomy to connectivity graphs. **a**, Three-dimensional rendering of the dendrites, axons and cell bodies of 14 neurons in the functionally imaged plane (coloured according to their orientation preference, key right, as in Fig. 1b), and the dendrites and cell bodies of all their postsynaptic targets traced in the EM volume (magenta, excitatory targets; cyan, inhibitory targets; spines on postsynaptic targets not shown; Supplementary Movie 5). Scale bar, 100 μ m. **b**, Directed network diagram of the functionally characterized cells and their targets, derived from **a**. Postsynaptic excitatory (magenta) and inhibitory (cyan) targets with cell bodies contained within the EM volume are drawn as circles. Other postsynaptic targets

(dendritic fragments) are drawn as squares. (From top to bottom and left to right: functionally characterized cells 5, 2, 7; 13, 6, 14; 1; 10; 11, 3; 9; 12, 4; and 8.) **c**, Three-dimensional rendering of the arbors and cell bodies of functionally characterized neurons, along with postsynaptic targets that either receive convergent input from multiple functionally characterized neurons, or were themselves functionally characterized (Supplementary Movie 5). **d**, A subset of the network graph showing only the connections in **c**, all independently verified (from top to bottom and left to right: functionally characterized cells 5, 2, 7; 13, 6; 10; 11, 3; 12, 9, 8 and 4).

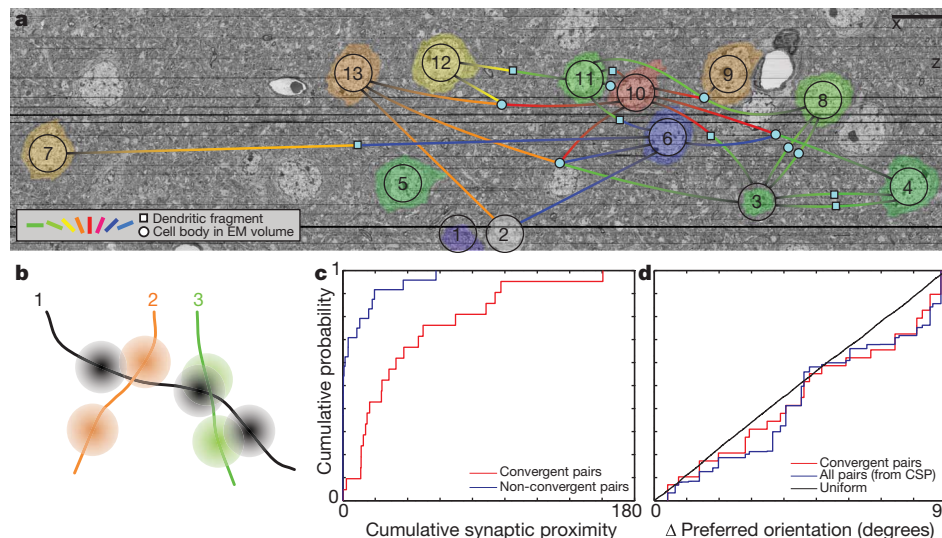


Figure 6 | Convergent synaptic input onto inhibitory interneurons is predicted by proximity, not function. **a**, A portion of the aligned and registered EM image series re-sliced parallel to the functionally imaged plane, through 1,153 EM sections. Overlaid is a network graph of the convergences onto inhibitory interneurons. Visually responsive cell bodies are pseudo-coloured according to their preferred orientation (as in Fig. 1b), and numbered as in Figs 4 and 5. Convergences onto inhibitory neuronal targets are represented by lines, corresponding to one or more synapses, leading either to filled cyan circles (targets traced to cell bodies in the EM volume) or squares (dendritic fragments). Cell 14 was partially contained in the EM volume and is not shown. Scale bars, 10 μ m. **b**, Diagram of cumulative synaptic proximity (CSP). Line segments represent axons, with three-dimensional Gaussians centred at each synapse. A CSP was calculated for each pair of orientation-

tuned neurons by summing all pair-wise overlaps of Gaussians from the two axons ($\sigma \approx 12 \mu$ m; Methods). **c**, Pairs of axons whose synaptic boutons were in close proximity were more likely to converge onto a common target. The CSP of axon pairs participating in convergences (red) was significantly greater than for non-converging pairs (blue; $P < 1.3 \times 10^{-5}$, two-sample Kolmogorov-Smirnov test, $n_{\text{convergent pairs}} = 21$, $n_{\text{non-convergent pairs}} = 24$). **d**, Convergences were not predicted by the difference in orientation preference between presynaptic cell pairs. The distribution of differences in orientation preference was not significantly different from a uniform distribution ($P > 0.30$, two-sample Kolmogorov-Smirnov test, $n_{\text{convergences}} = 29$) or a model distribution (Methods) based on CSP ($P > 0.68$, two-sample Kolmogorov-Smirnov test, $n_{\text{convergences}} = 29$).

were onto inhibitory targets, despite the preponderance of excitatory neurons in the cortex^{39–41}, and despite reports that 10–20% of the synapses made by pyramidal cells are onto inhibitory targets in cat¹⁷ and macaque⁴². Whether the higher percentage we observed is due to a species difference (ref. 43, compare ref. 44), or to the fact that we sampled synapses from proximal portions of the pyramidal cell axonal arbors, it resulted in our ability to sample a large number of convergences onto inhibitory targets (Fig. 5d).

Although the volume we imaged using EM was comparatively large, it proved to be near the minimum required to perform an analysis relating cortical function to network anatomy. We collected the series of wide-field ($350 \times 450 \mu$ m) high-resolution EM images to encompass the axonal arbors of the functionally characterized neurons and the dendrites of their targets. Nonetheless, we were limited by the shortest dimension in our volume (52 μ m), determined by the number of thin sections, so we could trace only 245 out of the thousands of synapses made by the functionally characterized neurons.

We anticipate that the size of serial EM volumes will increase substantially in the near future, owing to increases in imaging throughput and series length made possible by automated techniques^{45,46}. The time required to trace connectivity between neurons is likely to remain a limiting factor, although semi-automated techniques have already achieved 10-fold increases in throughput over purely manual approaches²². For large-scale reconstruction, data quality is paramount. To trace unlabelled, fine-calibre axons, it is essential to have minimum section thickness, minimum section loss, and optimal tissue quality.

In the current study, we found a large number of convergent inputs onto inhibitory neurons principally because they were densely innervated by the excitatory axons we reconstructed. Probing the network anatomy of more sparsely interconnected (and possibly weakly biased⁴⁷) excitatory neurons⁴⁸, however, will require larger samples. Here we sought to limit tissue damage from the infrared laser, so

two-photon calcium imaging was confined to a single plane, or less than 1% of the cells in the volume (Supplementary Fig. 8a). Recent advances in calcium imaging^{49,50}, however, should now allow physiological data to be collected from many more cells in a volume while maintaining tissue quality.

It is fortunate that increases in the dimensions of an EM-imaged volume, and the number of physiologically characterized cells within it, produce combinatorial increases in the number of network motifs⁸ that can be analysed in a single experiment (Supplementary Fig. 8b–d). In particular, if a population of neurons is sparsely sampled, the number of interconnections found between them increases as the square of the sampling density. With moderate gains in the number of functionally imaged cells, or in the volumes encompassed by EM reconstructions, insight into the functional logic of cortical networks should therefore increase at an accelerating pace.

METHODS SUMMARY

We performed two-photon imaging in the mouse visual cortex as described previously^{27,30} by recording calcium responses to visual stimuli consisting of drifting gratings in each of 16 directions. We then acquired an *in vivo* fluorescent anatomical volume after injecting the tail vein with SR101 (100 mM) to label vasculature. The animal was perfused transcardially (2% paraformaldehyde/2.5% glutaraldehyde) and the brain was processed for serial-section TEM. Serial thin (<50 nm) sections were cut, picked up on Pioloform-coated slot grids, and then post-stained with uranyl acetate and lead citrate. 1,215 serial sections were imaged at 120 kV on a JEOL 1200 EX with a custom scintillator atop optical-quality leaded vacuum glass at the end of a custom-built vacuum chamber extension. Custom software controlled automated *x*–*y* stage motion and image acquisition with a 2×2 array of CCD cameras (Imperx IPX-11M5) and Zeiss lenses. Images suitable for circuit reconstruction were acquired at a net rate of 5–8 megapixels per second. Camera images were aligned in two dimensions by registering adjacent camera images and dewarping, followed by histogram equalization and stitching. Then adjacent sections were registered and three-dimensional deformations were equalized in aligning the EM volume. Axonal and dendritic arbors of the functionally characterized neurons were manually reconstructed using

TrakEM2 and objects were classified using classical criteria³³. Neurons or dendritic fragments receiving synapses from multiple functionally characterized cells were included in analysis of convergence. For each synapse participating in a convergence, a second individual (blind to the original reconstruction) traced the pre- and post-synaptic processes, starting from the synapse. Segmentation that diverged between the two tracers was excluded from further analysis. Cumulative synaptic proximity (CSP) of pairs of axons was calculated by centring a three-dimensional Gaussian density function at each synaptic bouton and taking the sum of their dot products over all pairs of synapses.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 October 2010; accepted 10 January 2011.

- Ramón y Cajal, S. *Textura del Sistema Nervioso del Hombre y de los Vertebrados* (Moya, 1904).
- Binzegger, T., Douglas, R. J. & Martin, K. A. A quantitative map of the circuit of cat primary visual cortex. *J. Neurosci.* **24**, 8441–8453 (2004).
- Stepanyants, A. *et al.* Local potential connectivity in cat primary visual cortex. *Cereb. Cortex* **18**, 13–28 (2008).
- Mason, A., Nicoll, A. & Stratford, K. Synaptic transmission between individual pyramidal neurons of the rat visual cortex in vitro. *J. Neurosci.* **11**, 72–84 (1991).
- Markram, H., Lubke, J., Frotscher, M., Roth, A. & Sakmann, B. Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol. (Lond.)* **500**, 409–440 (1997).
- Thomson, A. M. & Bannister, A. P. Interlaminar connections in the neocortex. *Cereb. Cortex* **13**, 5–14 (2003).
- Braitenberg, V. & Schüz, A. *Cortex: Statistics and Geometry of Neuronal Connectivity* 2nd edn (Springer, Berlin, 1998).
- Song, S., Sjostrom, P. J., Reigl, M., Nelson, S. & Chklovskii, D. B. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* **3**, e68 (2005).
- Yoshimura, Y. & Callaway, E. M. Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. *Nature Neurosci.* **8**, 1552–1559 (2005).
- Yoshimura, Y., Dantzker, J. L. & Callaway, E. M. Excitatory cortical neurons form fine-scale functional networks. *Nature* **433**, 868–873 (2005).
- Heuser, J. E. *et al.* Synaptic vesicle exocytosis captured by quick freezing and correlated with quantal transmitter release. *J. Cell Biol.* **81**, 275–300 (1979).
- White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond. B* **314**, 1–340 (1986).
- Sorra, K. E. & Harris, K. M. Stability in synapse number and size at 2 hr after long-term potentiation in hippocampal area CA1. *J. Neurosci.* **18**, 658–671 (1998).
- Dacheux, R. F. & Raviola, E. The rod pathway in the rabbit retina: a depolarizing bipolar and amacrine cell. *J. Neurosci.* **6**, 331–345 (1986).
- Sterling, P. Microcircuitry of the cat retina. *Annu. Rev. Neurosci.* **6**, 149–185 (1983).
- Hamos, J. E., Van Horn, S. C., Raczkowski, D. & Sherman, S. M. Synaptic circuits involving an individual retinogeniculate axon in the cat. *J. Comp. Neurol.* **259**, 165–192 (1987).
- Kisvárdy, Z. F. *et al.* Synaptic targets of HRP-filled layer III pyramidal cells in the cat striate cortex. *Exp. Brain Res.* **64**, 541–552 (1986).
- Anderson, J. C., Douglas, R. J., Martin, K. A. & Nelson, J. C. Map of the synapses formed with the dendrites of spiny stellate neurons of cat visual cortex. *J. Comp. Neurol.* **341**, 25–38 (1994).
- Ahmed, B., Anderson, J. C., Martin, K. A. & Nelson, J. C. Map of the synapses onto layer 4 basket cells of the primary visual cortex of the cat. *J. Comp. Neurol.* **380**, 230–242 (1997).
- Tamas, G., Somogyi, P. & Buhl, E. H. Differentially interconnected networks of GABAergic interneurons in the visual cortex of the cat. *J. Neurosci.* **18**, 4255–4270 (1998).
- Shepherd, G. M. & Harris, K. M. Three-dimensional structure and composition of CA3→CA1 axons in rat hippocampal slices: implications for presynaptic connectivity and compartmentalization. *J. Neurosci.* **18**, 8300–8310 (1998).
- Mishchenko, Y. *et al.* Ultrastructural analysis of hippocampal neuropil from the connectomics perspective. *Neuron* **67**, 1009–1020 (2010).
- Anderson, J. R. *et al.* A computational framework for ultrastructural mapping of neural circuitry. *PLoS Biol.* **7**, e1000074 (2009).
- Sohya, K., Kameyama, K., Yanagawa, Y., Obata, K. & Tsumoto, T. GABAergic neurons are less selective to stimulus orientation than excitatory neurons in layer II/III of visual cortex, as revealed by *in vivo* functional Ca²⁺ imaging in transgenic mice. *J. Neurosci.* **27**, 2145–2149 (2007).
- Niell, C. M. & Stryker, M. P. Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* **28**, 7520–7536 (2008).
- Liu, B. H. *et al.* Visual receptive field structure of cortical inhibitory neurons revealed by two-photon imaging guided recording. *J. Neurosci.* **29**, 10520–10532 (2009).
- Kerlin, A. M., Andermann, M. L., Berezovskii, V. K. & Reid, R. C. Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron* **67**, 858–871 (2010).
- Runyan, C. A. *et al.* Response features of parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. *Neuron* **67**, 847–857 (2010).
- Ma, W. P. *et al.* Visual representations by cortical somatostatin inhibitory neurons — selective but with weak and delayed responses. *J. Neurosci.* **30**, 14371–14379 (2010).
- Ohki, K., Chung, S., Ch'ng, Y. H., Kara, P. & Reid, R. C. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* **433**, 597–603 (2005).
- Stosiek, C., Garaschuk, O., Holthoff, K. & Konnerth, A. *In vivo* two-photon calcium imaging of neuronal networks. *Proc. Natl Acad. Sci. USA* **100**, 7319–7324 (2003).
- Harris, K. M. *et al.* Uniform serial sectioning for transmission electron microscopy. *J. Neurosci.* **26**, 12101–12103 (2006).
- Peters, A., Palay, S. L. & Webster, H. D. *The Fine Structure of the Nervous System: Neurons and Their Supporting Cells* 3rd edn (Oxford Univ. Press, 1991).
- Alitto, H. J. & Dan, Y. Function of inhibition in visual cortical processing. *Curr. Opin. Neurobiol.* **20**, 340–346 (2010).
- Liu, B. H. *et al.* Intervening inhibition underlies simple-cell receptive field structure in visual cortex. *Nature Neurosci.* **13**, 89–96 (2010).
- Reynolds, J. H. & Heeger, D. J. The normalization model of attention. *Neuron* **61**, 168–185 (2009).
- Helmstaedter, M., Briggman, K. L. & Denk, W. 3D structural imaging of the brain with photons and electrons. *Curr. Opin. Neurobiol.* **18**, 633–641 (2008).
- Turaga, S. C. *et al.* Convolutional networks can learn to generate affinity graphs for image segmentation. *Neural Comput.* **22**, 511–538 (2010).
- Peters, A. & Kara, D. A. The neuronal composition of area 17 of rat visual cortex. II. The nonpyramidal cells. *J. Comp. Neurol.* **234**, 242–263 (1985).
- White, E. L. & Keller, A. *Cortical Circuits: Synaptic Organization of the Cerebral Cortex — Structure, Function and Theory* (Birkhauser, 1989).
- Markram, H. *et al.* Interneurons of the neocortical inhibitory system. *Nature Rev. Neurosci.* **5**, 793–807 (2004).
- McGuire, B. A., Gilbert, C. D., Rivlin, P. K. & Wiesel, T. N. Targets of horizontal connections in macaque primary visual cortex. *J. Comp. Neurol.* **305**, 370–392 (1991).
- Holmgren, C., Harkany, T., Svennenfors, B. & Zilberter, Y. Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *J. Physiol. (Lond.)* **551**, 139–153 (2003).
- Thomson, A. M., West, D. C., Wang, Y. & Bannister, A. P. Synaptic connections and small circuits involving excitatory and inhibitory neurons in layers 2–5 of adult rat and cat neocortex: triple intracellular recordings and biocytin labelling *in vitro*. *Cereb. Cortex* **12**, 936–953 (2002).
- Denk, W. & Horstmann, H. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biol.* **2**, e329 (2004).
- Hayworth, K., Kasthuri, N., Schalek, R. & Lichtman, J. Automating the collection of ultrathin serial sections for large volume TEM reconstructions. *Microsc. Microanal.* **12**, 86–87 (2006).
- Jia, H., Rochefort, N. L., Chen, X. & Konnerth, A. Dendritic organization of sensory input to cortical neurons *in vivo*. *Nature* **464**, 1307–1312 (2010).
- Ohki, K. & Reid, R. C. Specificity and randomness in the visual cortex. *Curr. Opin. Neurobiol.* **17**, 401–407 (2007).
- Grewe, B. F. & Helmchen, F. Optical probing of neuronal ensemble activity. *Curr. Opin. Neurobiol.* **19**, 520–529 (2009).
- Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6**, 875–881 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Raviola for discussions and technical advice on all aspects of EM; J. Stiles for support and advice concerning the alignment and stitching effort at The Pittsburgh Supercomputing Center; J. Lichtman for discussions and, along with K. Blum and J. Sanes, support from the Center for Brain Science; A. Cardona for help with TrakEM2, and many modifications of it; and H. Pfister and S. Warfield for discussions and help with computational issues. We also thank S. Butterfield for programming, and L. Benecchi and Harvard Medical School EM Core Facility for technical support. For technical help with EM and its interpretation, we thank K. Harris, M. Bickford, M. Ellisman, K. Martin and T. Reese. We thank J. Assad, R. Born, J. Maunsell, J. Stiles, E. Raviola, and members of the Reid laboratory for critical reading of the manuscript. This work was supported by the Center for Brain Science at Harvard University, Microsoft Research, and the NIH through the NEI to R.C.R. (EY10115 and EY18742) and through resources provided by the NRBSC (P41 RR06009), which is part of The Pittsburgh Supercomputing Center; and by fellowships from Harvard Center of Neurodegeneration and Repair to D.D.B. and the NEI to W.-C.A.L. (EY18532).

Author Contributions D.D.B., E.R.S., S.Y. and R.C.R. designed and built the TEMCA system, and D.D.B. programmed it. A.M.K. and M.L.A. performed the *in vivo* imaging and A.M.K. analysed it. D.D.B. processed the tissue, and D.D.B. and W.-C.A.L. aligned the block with the *in vivo* imaging. W.-C.A.L. sectioned the series, and D.D.B., W.-C.A.L. and H.S.K. imaged it on the TEMCA. G.H. and A.W.W. stitched and aligned the images into a volume. D.D.B., W.-C.A.L. and H.S.K. did most of the segmentation. D.D.B., W.-C.A.L. and S.Y. performed quantitative analysis on the tracing. D.D.B., W.-C.A.L. and R.C.R. designed the experiment and wrote the paper, with considerable help from the other authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.C.R. (clay_reid@hms.harvard.edu).

METHODS

Animal preparation and *in vivo* calcium imaging. We imaged the visual cortex of an adult mouse (male *Thy1-YFP-H*⁵¹), as described previously^{27,30}, using a custom-built two-photon microscope with a 20× objective (Olympus, 0.95 NA) at a frame rate of 2 Hz. Randomized drifting gratings (100% contrast, 2 Hz) were presented for 4 s, followed by 4 s of uniform mean luminance on an LCD monitor, positioned 18 cm from the contralateral eye, spanning approximately 80° (azimuth) by 60° (elevation) of visual space. To determine orientation tuning, square-wave gratings (0.05 cycles per degree) were presented at 16 directions of motion. Cellular responses were measured with the calcium indicator Oregon Green BAPTA-1 AM (OGB; Invitrogen). We next injected the tail vein with SR101 (100 mM; Invitrogen) to fluorescently label vasculature and acquired an *in vivo* fluorescent anatomy volume of the region (600 × 600 × 250 μm) surrounding the calcium-imaged plane (Fig. 1c).

EM material preparation. Following *in vivo* two-photon imaging the animal was perfused transcardially (2% paraformaldehyde/2.5% glutaraldehyde) and the brain was processed for serial-section TEM. Seven 300-μm-thick coronal vibratome sections were cut. The vibratome sections were post-fixed and stained with 1% osmium tetroxide/1.5% potassium ferrocyanide followed by 1% uranyl acetate, dehydrated with a graded ethanol series, and embedded in resin (TAAB 812 Epon, Canemco). After locating the calcium-imaged region by matching vasculature between *in vivo* fluorescence and serial thick (1 μm) toluidine blue (EMS) sections cut from an adjacent vibratome section, 4,150 serial thin (<50 nm) sections were cut on an ultramicrotome (Leica UC6) using a 35° diamond knife (EMS-Diatome) and picked up on 1 × 2 mm dot-notch slot grids (Synpatek) that were coated with a pale gold Pioloform (Ted Pella) support film, carbon coated, and glow-discharged. Following section pickup, grids were post-stained with uranyl acetate and lead citrate. Of the 4,150 serial sections, we targeted 1,215 serial sections for imaging that contained the greatest diversity of orientation selectivities and the fewest lost sections. Within this range, a total of 28 sections were lost: 2 during pickup, 16 during post-staining, 2 before imaging, and 8 during imaging.

Transmission electron microscope camera array (TEMCA). A 120 kV TEM (JEOL 1200 EX) was placed atop scaffolding with a custom-built 1.2 m vacuum chamber extension beneath the microscope's final projector lens, culminating in a 226 mm diameter scintillator (Grant Scientific) resting above optical-quality leaded vacuum glass (Computer Optics) (Supplementary Fig. 2). Custom software (LabView) controlled x-y stage motion and image acquisition from a 2 × 2 array⁵² of CCD cameras (Imperx IPX-11M5) coupled to lenses (Zeiss Makroplanar T* 50 mm F/2.0 ZF). Each acquired camera frame was cross-correlated, translated, and summed by the camera's dedicated acquisition computer, which wrote data over a local area network to storage servers (InterPRO Microsystems). Magnification at the microscope was 2,000×, accelerating potential was 120 kV, and beam current was ~90 μA through a tungsten filament. Images suitable for circuit reconstruction were acquired at a net rate of 5–8 megapixels per second.

EM image registration and correspondence. A detailed description of the computational methods for registration and alignment of the EM volume will appear elsewhere. Briefly, camera images were aligned in two dimensions by registering adjacent camera images and dewarping, followed by histogram equalization and stitching. Then adjacent sections were registered and three-dimensional bending stresses were equalized to align the EM volume. The imaged volume will be publicly accessible at the Whole Brain Catalog (<http://wholebraincatalog.org>) and the UCSD Cell Centered Database⁵³ (<http://ccdb.ucsd.edu/>, accession number 8448).

The aligned volume (sections 2,990–4,150) was imported into a customized version of TrakEM2⁵⁴ for visualization and segmentation. Although we targeted imaging to 1,215 serial-sections, the final stitched and registered data set contains 1,153 sections, since some fragmentary material, typically small pieces cut from an edge of the EM block face, was excluded. Volume renderings of the raw, unsegmented EM image data (Fig. 2b, e) were made in TrakEM2 using a false-colour lookup table. To find the correspondence between the cells imaged *in vivo* with those in the EM data set, we used successively finer scales of vasculature and then somata to re-locate the calcium-imaged neurons in the EM-imaged volume (Fig. 3 and Supplementary Fig. 3). Pixel size (4 nm per pixel) was calibrated using a latex bead-coated replica grating (EMS). Section thicknesses of ~45 nm were estimated using the cylindrical diameters method⁵⁵.

Tracing and independent verification. Axonal and dendritic arbors of the functionally characterized neurons were traced by manually placing a series of marker points down the midline of each process and rendering a wire-frame model of the arbors using TrakEM2⁵⁴. Infrequently, ambiguities arose from poor local data quality or missing sections. In this case, tracing of a process was halted. Fine axons generally could not be traced across three or more consecutive missing sections. Synapses were identified using classical criteria³³. For each synapse on the axon of

a functionally characterized cell, we traced the dendrites of the postsynaptic neuron either to the boundaries of the volume or centripetally back to the cell body. Neurons or dendritic fragments receiving synapses from multiple functionally characterized cells were considered convergence targets.

A second individual who had not previously reconstructed the pre- or postsynaptic process, and was blind to previous tracing work, independently verified the anatomical connectivity underlying convergences and connections between functionally characterized cells. Verification started at each synapse (small yellow spheres, Fig. 4a and Supplementary Fig. 7) and the pre- and postsynaptic processes were each traced centripetally (thick tracing, Fig. 4a and Supplementary Fig. 7). When the initial reconstruction and the subsequent verification of the reconstruction diverged, that connection and all segmentation work distal from the point of the divergence was excluded from further analysis. The verification process revealed that our initial segmentation effort was highly reproducible, consistent with earlier tests looking at reproducibility of full axonal arbors (Supplementary Fig. 6). All postsynaptic dendritic processes were verified. Of 64 presynaptic axonal contacts participating in convergences, 55 were independently verified. Three experienced team members did the bulk of the tracing and verification over a period of three months. Apical dendrites and dendritic spines (Fig. 5a, c and Supplementary Figs 3, 7) not participating in either the full (Fig. 5b) or independently verified (Fig. 5d) connectivity graphs were traced by a small team of students and were not included in the analysis.

Categorization of postsynaptic targets as inhibitory or excitatory. Dendrites postsynaptic to axons of physiologically characterized cells were unambiguously identified as belonging to inhibitory or excitatory neurons. Inhibitory dendrites were smooth and densely coated with asymmetric shaft synapses, whereas excitatory dendrites were spiny and had fewer shaft synapses (Supplementary Fig. 4a, b). One-hundred-and-thirty-three (54%) of the synapses were onto dendritic spines, and 112 (46%) were onto dendritic shafts. Only four (1.6%) of the shaft synapses were onto excitatory targets, and 12 (4.9%) of the spine synapses were onto inhibitory targets. Starting at 10 randomly selected synapses made by functionally characterized neurons, the postsynaptic dendrites' percentage of synapses on spines was calculated by tracing along the dendrite and counting up to 10 synapses (including those on spines and the shaft) in either direction. The distribution of spine synapses and shaft synapses was non-overlapping between cell types⁴² (Supplementary Fig. 4b) and was consistent with the prior categorization of the target. In all cases ($n = 8$) where the postsynaptic dendrite was categorized as inhibitory and could be traced to a cell body (cyan circles in Figs 5, 6), asymmetric synapses were found contacting the soma (Supplementary Fig. 4c). Cell 2, which was visually responsive, but not selective for stimulus orientation (Supplementary Fig. 1), exhibited all of the aforementioned characteristics. Furthermore, its axon made symmetric contacts with both dendrites and somata (Supplementary Fig. 4d).

Data analysis. Calcium-imaging data analyses were performed with Matlab (MathWorks) and ImageJ (NIH) as described previously^{27,30}. Only neurons with an estimated photonic noise floor <3% $\Delta F/F$ were included. Neurons were considered visually responsive if the response to the best direction was >6% $\Delta F/F$ (morphologically identified pyramidal neurons) or >4% $\Delta F/F$ (non-pyramidal neurons). Orientation tuning curves for each neuron were generated from the average responses to 16 directions fit to the sum of two Gaussians (Supplementary Fig. 1) and the preferred orientation was estimated from this fit. Orientation selectivity was calculated as the magnitude of the vector average divided by the sum of all responses: $(\sum R(\theta_i) \sin(2\theta_i))^2 + (\sum R(\theta_i) \cos(2\theta_i))^2)^{1/2} / \sum R(\theta_i)$ (refs 56, 57).

The cumulative synaptic proximity (CSP) for a pair of axons was calculated by centring a clipped three-dimensional Gaussian density function at each synaptic bouton, calculating the dot product of these Gaussians for each pair of boutons (one bouton from each axon), and summing over all pairs (Fig. 6b). The dot products were normalized so that two superimposed synapses yielded a value of 1.0. The distribution of CSPs between convergent and non-convergent pairs was significantly different (two-sample Kolmogorov-Smirnov test, $\sigma = \sim 12 \mu\text{m}$). A model population of convergences was generated by allowing each possible pair of physiologically characterized cells to contribute a number of convergences to the model proportional to the pair's CSP, with high CSP pairs contributing many convergences, and low CSP pairs contributing few. The distribution of relative preferred orientations from the model population (Fig. 6d, blue line) did not differ significantly from either the uniform distribution or the population of convergent pairs (two-sample Kolmogorov-Smirnov test). A separate set of model populations was generated in which the probability of convergence by a cell pair increased with the similarity of their orientation preferences. Drawing randomly from these model populations we found our sample size (Fig. 5d, $n_{\text{convergences}} = 29$) was sufficient to detect a statistically significant difference from the uniform distribution (two-sample Kolmogorov-Smirnov test, $\alpha = 0.05$, power ~0.80) in strongly biased model populations, that is, populations in which convergent pairs with orthogonal preferred orientation tunings were excluded.

However, in hypothetical populations exhibiting weak biases larger sample sizes are necessary (data not shown).

51. Feng, G. *et al.* Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. *Neuron* **28**, 41–51 (2000).
52. Peltier, S. *et al.* Design of a new 8k x 8k lens coupled detector for wide-field, high-resolution transmission electron microscopy. *Microsc. Microanal.* **11**, 610–611 (2005).
53. Martone, M. E. *et al.* A cell-centered database for electron tomographic data. *J. Struct. Biol.* **138**, 145–155 (2002).
54. Cardona, A. *et al.* An integrated micro- and macroarchitectural analysis of the *Drosophila* brain by computer-assisted serial section electron microscopy. *PLoS Biol.* **8**, e1000502 (2010).
55. Fiala, J. C. & Harris, K. M. Cylindrical diameters method for calibrating section thickness in serial electron microscopy. *J. Microsc.* **202**, 468–472 (2001).
56. Ringach, D. L., Shapley, R. M. & Hawken, M. J. Orientation selectivity in macaque V1: diversity and laminar dependence. *J. Neurosci.* **22**, 5639–5651 (2002).
57. Worgotter, F. & Eysel, U. T. Quantitative determination of orientational and directional components in the response of visual cortical cells to moving stimuli. *Biol. Cybern.* **57**, 349–355 (1987).

Wiring specificity in the direction-selectivity circuit of the retina

Kevin L. Briggman¹, Moritz Helmstaedter¹ & Winfried Denk¹

The proper connectivity between neurons is essential for the implementation of the algorithms used in neural computations, such as the detection of directed motion by the retina. The analysis of neuronal connectivity is possible with electron microscopy, but technological limitations have impeded the acquisition of high-resolution data on a large enough scale. Here we show, using serial block-face electron microscopy and two-photon calcium imaging, that the dendrites of mouse starburst amacrine cells make highly specific synapses with direction-selective ganglion cells depending on the ganglion cell's preferred direction. Our findings indicate that a structural (wiring) asymmetry contributes to the computation of direction selectivity. The nature of this asymmetry supports some models of direction selectivity and rules out others. It also puts constraints on the developmental mechanisms behind the formation of synaptic connections. Our study demonstrates how otherwise intractable neurobiological questions can be addressed by combining functional imaging with the analysis of neuronal connectivity using large-scale electron microscopy.

The computation of motion direction by direction-selective retinal ganglion cells (DSGCs), discovered almost 50 years ago¹, has defied comprehensive explanation, partly because the wiring diagram of the neuronal circuit underlying this computation is still not known in sufficient detail. DSGCs respond strongly to motion oriented along a preferred direction, but not to null-direction (where null direction is 180° from the preferred direction) motion. In theory, this asymmetry could arise from increased inhibition during null-direction motion, increased excitation during preferred-direction motion, or through a combination of both mechanisms. The asymmetry could be implemented at the structural level, in the wiring of the direction-selectivity circuitry, or could result from unequal synaptic strengths in an otherwise structurally symmetric circuit. Barlow and Levick² favoured a mechanism involving the selective pre-emption of responses during null-direction motion by lateral inhibition. The magnitude of the inhibitory synaptic input to DSGCs is indeed spatially asymmetric, as patch-clamp recordings from DSGCs have shown^{3,4}. The main source of this inhibition is starburst amacrine cells (SACs), retinal interneurons^{5,6} that are necessary in the direction-selectivity circuit⁷ and release both GABA (γ -aminobutyric acid) and acetylcholine⁸. Furthermore, SAC dendrites have been shown to be direction-selective, preferring centrifugal motion⁹.

Asymmetric connectivity between SACs and DSGCs forms the basis of most models for how direction selectivity is computed^{4,10–15}. Dual whole-cell recordings have shown inhibitory connections between SACs and DSGCs to be stronger when the SAC's soma is located on the null side (that is, the side from which null-direction stimuli approach) of the DSGC's soma^{4,16,17}. However, it is not known whether individual synapses from null-side SACs have a higher unitary conductance^{16–18}, for example, as a result of developmental strengthening¹⁹, or, alternatively, whether this functional asymmetry is due to an increased number of synaptic contacts. It is also unknown whether synapse formation or strengthening is dependent mainly on the relative location of SAC and DSGC somata or rather on the direction of the SAC dendritic branch. The latter would be expected if each DSGC optimally collects appropriately selective input signals from SAC

dendrites. Both questions are difficult to address solely by physiological means. Rather, structural information about the location of individual SAC-to-DSGC synaptic contacts and the direction of the pre-synaptic SAC dendrites is needed.

Attempts to study direction-selectivity circuit anatomy by light microscopy have led to contradictory results, with some reports⁴ showing a higher number of neurite proximities from null-side SACs and others not^{20–22}. Electron microscopy studies of the direction-selectivity circuit^{5,23} have, on the other hand, not been able to reconstruct large fractions of SACs and DSGCs in the same piece of tissue. This illustrates the general difficulty of acquiring and analysing three-dimensional ultrastructural data to determine neural circuit diagrams. Such data need to have sufficient resolution to reliably follow small calibre (as small as a few tens of nanometres) neuronal processes and, at the same time, need to cover the spatial distances over which individual neurons project, typically at least some hundreds of micrometres.

Serial-section electron microscopy (SSEM), which was used for the reconstruction of the *Caenorhabditis elegans* nervous system²⁴, is laborious, uses thicker sections and data are occasionally corrupted by sectioning and imaging artefacts. These problems are ameliorated by automating the acquisition of EM data, as in serial block-face electron microscopy (SBEM)²⁵. SBEM provides the necessary three-dimensional resolution and field of view to follow thin neurites across hundreds of micrometres of complex neuropil.

What complicates the purely anatomical analysis of the direction-selectivity circuit is that a DSGC's preferred direction cannot be deduced from its dendritic morphology^{26,27}. We, therefore, combined two-photon functional calcium imaging in the intact retina²⁸ with a SBEM-based reconstruction of the SAC–DSGC circuitry in the same piece of tissue. We found a strong asymmetry in the number of synaptic contacts, most of which are from SAC dendrites that are roughly oriented along the null direction of DSGCs, supporting a mechanism of null-direction inhibition. We also found evidence that the orientation of an individual SAC dendrite rather than the axis between the SAC and DSGC somata determines whether or not a synapse forms.

¹Max Planck Institute for Medical Research, Department of Biomedical Optics, Heidelberg 69120, Germany.

Functional and structural identification of DSGCs

To identify the preferred directions of On–Off DSGCs, we labelled the ganglion cell layer of an adult mouse retina by bulk electroporation⁴⁸ with the membrane-impermeable form of Oregon Green 488 BAPTA-1, a calcium indicator. This avoids the damage that would inevitably result from the pipette penetration needed for acetoxymethyl ester-based loading^{29,30} and would possibly result from exposing the retina to the detergents used during this procedure. We then used two-photon-excited fluorescence imaging³¹ to characterize the response properties of ganglion cells while projecting moving-bar stimuli (oriented in eight equally spaced directions) onto the photoreceptors³². We imaged 634 neuronal somata in a 300 μm by 300 μm large region of the ganglion cell layer (Fig. 1b). Among those were 25 On–Off DSGCs with preferred directions clustering in 4 groups (Fig. 1a and Supplementary Fig. 1). We denoted those groups, which are known to correspond to the cardinal visual axes³³, as northward (N), eastward (E), southward (S) and westward (W). The cells (6 N, 8 E, 7 S, and 4 W) were arranged in a mosaic pattern (Fig. 1b). Immediately after two-photon imaging, we fixed and stained the retina (see Methods) and prepared it for SBEM. To assist traceability, the tissue was specially treated to preferentially label cell surfaces and to leave intracellular structures unstained. The acquired SBEM volume was 350 \times 300 \times 60 μm^3 in extent, spanned the inner plexiform layer, and contained the ganglion cell layer and part of the inner nuclear layer. The lateral resolution was 16.5 nm \times 16.5 nm and the section thickness (z-resolution) was 23 nm. All calcium-imaged somata (Fig. 1c) were included in the acquired SBEM volume.

Vasculature landmarks were used to identify the somata of the recorded DSGCs in the SBEM volume (Fig. 1c). Beginning at their somata, we traced the dendritic trees of six DSGCs (Fig. 2a; 2 N, 1 E, 1 S and 2 W cells). Instead of contouring each dendrite, we traced skeletons along the centre lines of the dendrites, which speeds up the tracing process considerably (M. Helmstaedter, K. L. Briggman

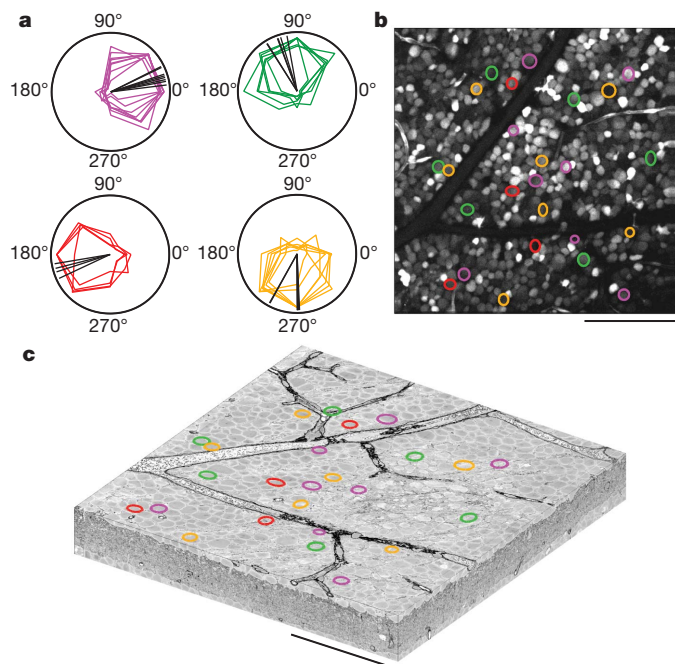


Figure 1 | Functional characterization of DSGCs and their localization within the SBEM volume. **a**, Polar tuning curves for 25 DSGCs sorted and colour-coded by preferred direction. Black lines indicate the direction of the vector-summed responses. **b**, **c**, The corresponding soma locations superimposed onto a two-photon image from the recorded region of the ganglion cell layer (**b**) and the acquired SBEM volume (**c**). Note the Y-shaped blood vessel visible in both **b** and **c**. Scale bars are 100 μm .

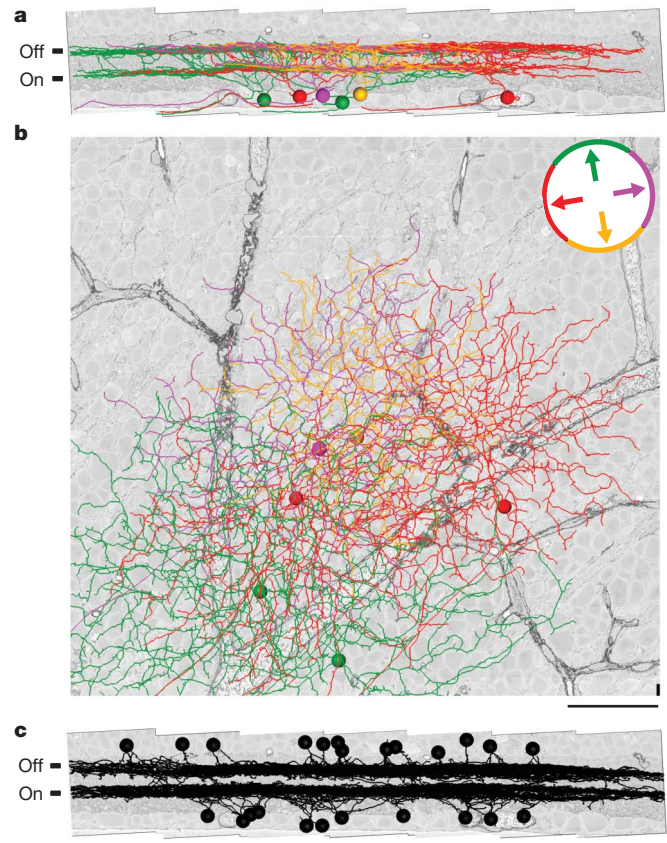


Figure 2 | Skeleton reconstructions of DSGCs and SACs. **a**, **b**, DSGCs, colour-coded by preferred direction (inset), projected parallel to (**a**) and normal to (**b**) the plane of the retina. Note bi-stratification in the inner plexiform layer. **c**, Parallel projections of 24 SACs (11 On SACs, 13 Off SACs, black). Scale bars are 50 μm .

and W. Denk, manuscript submitted). The resulting dendritic trees all ramified in two distinct sublayers in the inner plexiform layer and overlapped each other horizontally (Fig. 2a, b). The output synapses of SACs are formed at varicosities along the distal third of their dendrites⁵ and are geometrically conspicuous, with the presynaptic varicosity wrapping around postsynaptic dendrites³⁴. We, therefore, identified such varicose contacts (Fig. 3a, $n = 24$ contacts) on both the On and Off dendrites of each of the DSGCs and traced the putative presynaptic neurites back to their respective somata. Starting at these somata, we then skeletonized most of their dendritic trees, which substantially overlapped the dendritic fields of the DSGCs (Fig. 2c and Supplementary Figs 2 and 7). In every instance, the back-traced cell was a SAC, recognizable by its radially symmetric morphology and co-stratification with either On or Off DSGC arborizations ($n = 11$ On, 13 Off SACs). Given an estimated SAC density (including On and Off SACs) of 2,000 mm^{-2} for the mouse³⁵, we skeletonized 11% of all SAC somata in the data set.

Synaptic connections between SACs and DSGCs

To identify all additional potential contacts between the 24 SACs and 6 DSGCs that were reconstructed, we next inspected all locations where a SAC and a DSGC skeleton came within 1.5 μm of each other. Of 9,260 such locations, 831 were varicose contacts and were marked as putative synapses (Fig. 3a). In 2,650 other cases, a non-varicose, thin part of a SAC dendrite touched a DSGC dendrite (Fig. 3c). Such configurations were marked as ‘incidental’ contacts and are expected owing to the tight co-fasciculation of SAC and DSGC dendrites⁵. We occasionally found a configuration (Fig. 3e) where a SAC varicosity and a DSGC dendrite came within tens of nanometres of each other

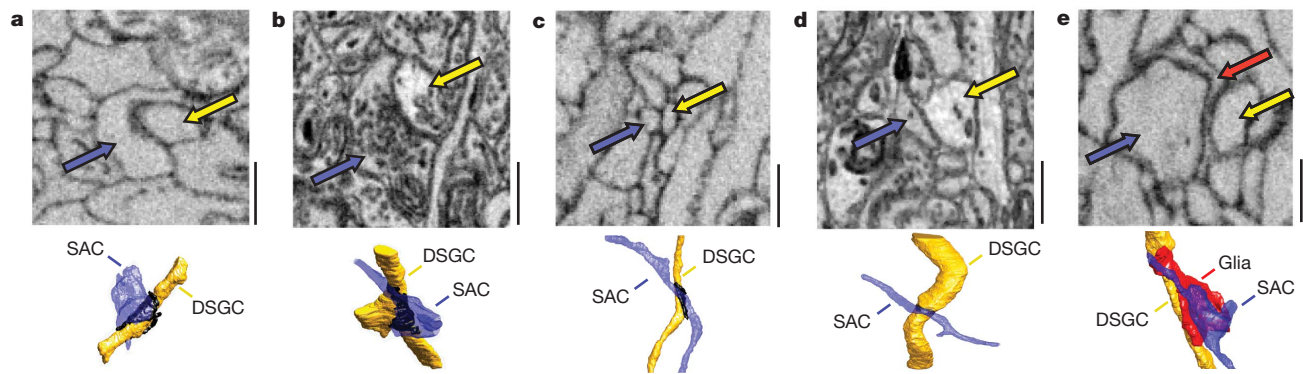


Figure 3 | Contact geometries. **a**, Typical putative synaptic contact in the surface-labelled sample in cross-section and as a three-dimensional rendering, with the contact area in black. **b**, Identified SAC-to-DSGC synapse in the conventionally stained SBEM volume. Note vesicles associated with the presynaptic surface and the dense staining of the contact. In both **a** and **b**, the

SAC varicosity (blue) wraps around the DSGC dendrite (yellow). **c**, **d**, Incidental contacts (thin regions of SAC and DSGC dendrites touching) in the surface-labelled (**c**) and the conventionally stained SBEM (**d**) volumes. **e**, SAC varicosity located very close to a DSGC, but separated by a thin Müller-cell glial process (red). Scale bars are 500 nm.

but did not make any direct contact and were instead separated by a thin, sheet-like process (probably from a Müller glia cell). Such appositions would probably be misconstrued as actual contacts using even the highest possible resolution in the (diffraction-limited) light microscope.

To confirm that the putative synapses are likely to be actual synapses, we examined varicose SAC–DSGC contacts in a different SBEM data set that was more conventionally stained (Fig. 3b). All varicose contacts inspected in this data set ($n = 43$) contained ultrastructural details typical of synapses. In addition, no synaptic specializations were seen at incidental SAC–DSGC contacts in this data set (Fig. 3d). The following analysis includes only the 831 contacts that were marked as putative synapses in the surface-stained data set (henceforth referred to as ‘synapses’).

We first examined the specificity of SAC–DSGC synapses from the perspective of individual SACs. We chose one Off and one On SAC (Fig. 4a and Supplementary Fig. 3) that each overlapped with the six DSGCs and colour-coded their output synapses by the preferred directions of the respective DSGCs (Fig. 4b; purple, E; green, N; red, W; orange, S). In addition, we identified all of the remaining varicosities on the dendrites of these two SACs (413 On SAC; 452 Off SAC; black dots in Fig. 4a and Supplementary Fig. 3). Output synapses preferred DSGCs with a preferred direction antiparallel to the SAC dendrite (and hence aligned with the null direction). For example, the northward oriented branches of the SACs mostly synapsed onto the southward preferring (orange) DSGC. Despite a large overlap of these northward branches with the dendritic trees of a westward (red) and eastward (purple) DSGC (Fig. 4a), they avoided synapsing onto them. The specificity is even more apparent in the outputs to the two westward (red) DSGCs. This pattern of specificity was found across all reconstructed SACs (Fig. 4b). A given SAC branch does not exclusively synapse onto only one type of DSGC; synapses onto DSGCs with different preferred directions sometimes occur, in particular for dendrites oriented in between the cardinal directions. We observed no obvious difference in the selectivity between On and Off sublayers (data not shown).

We next examined the specificity of synapses from the perspective of individual DSGCs. For each SAC–DSGC synapse, we constructed a vector oriented from its presynaptic SAC soma to the synapse location (Fig. 5a and Supplementary Fig. 5). We measured the angle (dendrite angle) between this vector and the 0° stimulus direction. The distribution of dendrite angles was strongly non-uniform for each DSGC (Fig. 5b), with most SAC dendrites oriented opposite to the DSGC’s preferred direction (Fig. 5c). The difference between dendrite angle and preferred direction was $165.2^\circ \pm 51.7^\circ$ (mean \pm s.d., $n = 831$). A

similar distribution of dendrite angles was observed in the conventionally stained retina data set (Supplementary Fig. 4), where, however, the preferred direction of the DSGC was not a priori known. Consistent with electrophysiological recordings, which found a higher

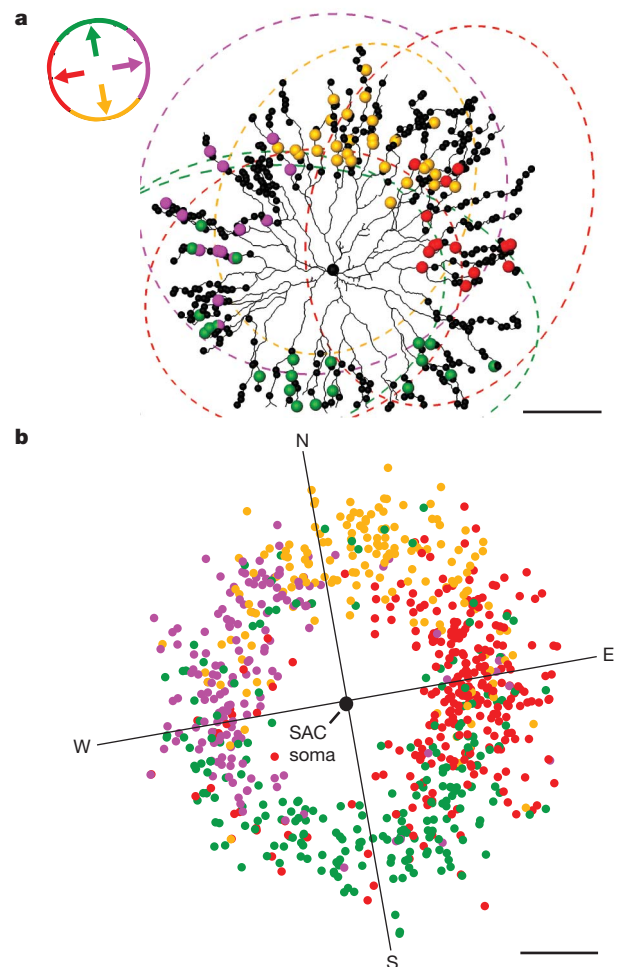


Figure 4 | Specificity of SAC outputs. **a**, An Off SAC (black skeleton), with varicosities indicated by black dots. DSGC dendritic trees are indicated by colour-coded dashed ellipses. Synapses are colour-coded by the preferred direction of the postsynaptic DSGC. **b**, Output synapse locations ($n = 831$ synapses) relative to SAC somata from all 24 SACs. Scale bars are 50 μ m.

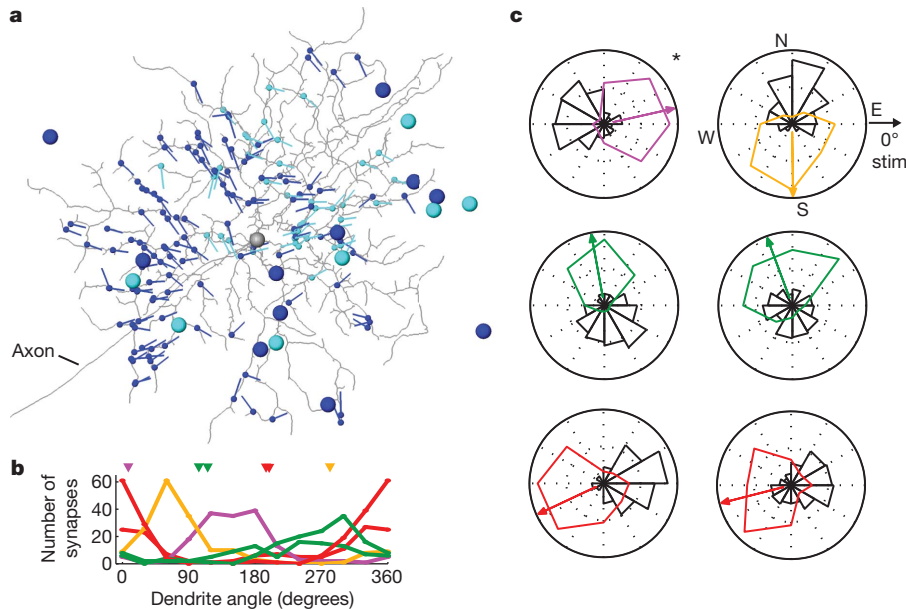


Figure 5 | Specificity of DSGC inputs. **a**, DSGC (grey skeleton) and the connected On and Off SAC somata (large cyan and blue circles, respectively) and associated SAC input synapses (smaller cyan and blue circles) from 18 SACs. **b**, The distribution of all SAC dendrite angles (θ_{dendr}) for each of the six DSGCs; θ_{dendr} is defined by the vectors (cyan and blue lines in **a**) oriented from

overall inhibitory conductance driven by SACs located on the null side^{4,16–18}, we found a larger number of synapses received from null-side versus preferred-side SAC somata (524 versus 41 synapses, where ND indicates null direction; $|\theta_{\text{soma}} - \theta_{\text{ND}}| < 45^\circ$ versus $|\theta_{\text{soma}} - \theta_{\text{ND}}| > 135^\circ$; Fig. 6b).

Although we found a strong correlation between SAC dendrite angles and DSGC null directions, we have not ruled out that the probability of forming a synapse between a SAC and a DSGC dendrite is solely determined by the relative locations of the corresponding somata, that is, by the angle between the null direction and the SAC-soma-DSGC-soma axis (Fig. 6a). In this case the connectivity should not depend on whether a dendrite is aligned more or less closely with the null direction than the soma-soma axis (Fig. 6a, upper panel); however, if, instead, the dendrite angle is the determinant then it should (Fig. 6a, lower panel). Our analysis (Fig. 6b) shows that for dendrite angles closer to the null direction ($\gamma < 0^\circ$) the actual connectivity is substantially higher. This is not caused by uneven sampling because the distribution of all contacts (incidental and varicose, Fig. 6c) is unbiased. We found the strongest bias when the soma-soma axis is between 45° and 135° off the null direction (Fig. 6b; $\gamma < 0^\circ$ for 75.2% of the synapses). We also calculated the mean of γ separately for different soma-soma angle ranges (Fig. 6d). In the range from 45° to 135° , the connected dendrites run $24.3^\circ \pm 2.8^\circ$ (mean \pm s.e.m.) closer to the null direction.

Discussion

Our data show that SAC dendrites selectively synapse with a DSGC if they are oriented along its null direction. This pattern provides the structural substrate for the functional asymmetry in the inhibitory input currents observed in DSGCs^{3,4}. The wiring specificity is apparent both from the perspective of the SACs' outputs (Fig. 4) and that of the DSGCs' inputs (Fig. 5). Dendritic branches of SACs are individually direction selective for centrifugal motion⁹, with several mechanisms likely to contribute^{36–38}. Our data support the view that DSGCs acquire their direction selectivity by predominately collecting those SAC inputs that suppress null-direction excitation, that is, from branches oriented along the null direction, and are consistent with

SAC somata to synapse location. Triangle markers indicate the preferred direction for each DSGC. **c**, Polar histograms of θ_{dendr} (black, plotted as the square root of θ_{dendr} frequencies) together with the DSGC tuning curves (as in Fig. 1). Asterisk denotes the DSGC shown in **a**.

the idea that null-direction SAC input inhibits the initiation of DSGC dendritic spikes during null-direction motion¹⁴.

The specificity of each SAC dendritic branch for selecting a post-synaptic target goes well beyond the notion that neuron A selectively

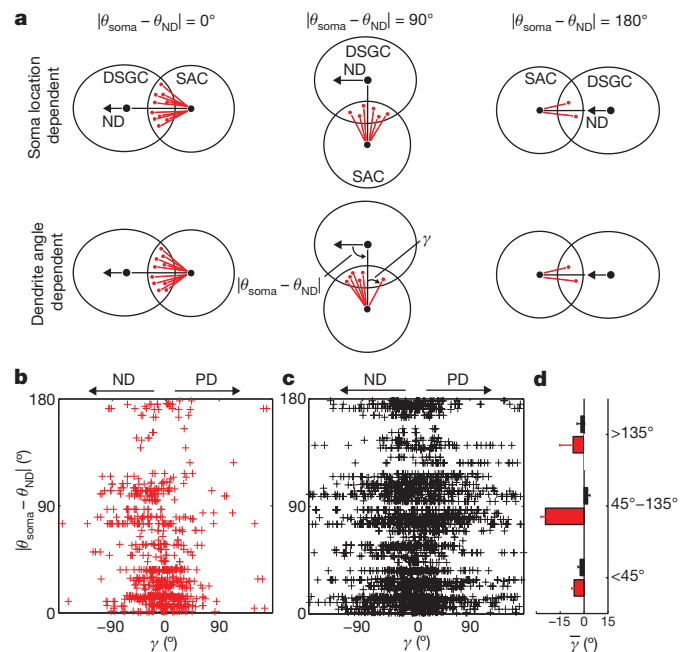


Figure 6 | Dendrite-angle distribution. **a**, Schematic synapse distributions expected for a mechanism depending only on soma angles (upper) and for a mechanism depending on dendrite angles (lower). The soma angle $|\theta_{\text{soma}} - \theta_{\text{ND}}|$ and relative dendrite angle $\gamma = (\theta_{\text{soma}} - \theta_{\text{dendr}}) \times \text{sign}(\theta_{\text{soma}} - \theta_{\text{ND}})$ are indicated, with $\gamma < 0$ indicating dendrite directions closer to the null direction than the soma-soma vector. **b**, Angular distribution of synapses ($n = 831$). **c**, Distribution of all synapses and incidental touches ($n = 3,481$). **d**, The mean \pm s.e.m. for γ in three different $(\theta_{\text{soma}} - \theta_{\text{ND}})$ ranges.

wires to neuron B, which is all that electrophysiological measurements can test. Instead the dendrite angle has an additional, perhaps dominant, role (Fig. 6), which is consistent with SAC dendrites acting as independent computational units⁹. This implies that the mechanism that determines whether an actual synapse is formed cannot depend on cell identities alone but uses the local dendrite geometry or activity.

SAC synapses onto DSGCs (Fig. 5) are not restricted to the null-side half of the DSGC dendritic tree but can be found across the entire dendritic field (Fig. 5a and Supplementary Fig. 5). Because SACs and DSGCs have similar dendritic-field diameters, some somata of strongly connected SACs lie near or actually on the preferred side of a DSGC's soma (Fig. 5a and Supplementary Fig. 5). This arrangement is consistent with the observation that DSGCs are direction-selective for local motion (that is, motion of a small spot within the dendritic field)², and also supports the notion that asymmetric null-direction lateral inhibition underlies the non-discriminating zone along the preferred-side edge of DSGC receptive fields^{2,39}.

We have presumed that the asymmetrically arranged SAC synapses are GABAergic, consistent with paired SAC–DSGC recordings^{4,16,17} and the finding that GABA antagonists block direction-selective responses in DSGCs^{40,41}. SACs also synthesize and release acetylcholine^{6,8,42}, but the role of acetylcholine in the direction-selectivity circuit remains unclear¹³. Some models^{11,15} argue that, in addition to asymmetric GABA release during null-direction motion, acetylcholine excites DSGCs asymmetrically during preferred-direction motion, leading to a SAC-generated ‘push–pull’ mechanism. Such a mechanism would, however, require contacts from preferred-direction-oriented dendrites, which are very sparse (Fig. 5). Further analysis of our EM data sets should reveal whether a different type of push–pull mechanism is present, whereby bipolar-cell terminals are inhibited²³ in a direction-specific manner.

Cholinergic currents have been recorded in DSGCs during both preferred-direction and null-direction motion⁴³. A study¹⁷ of SAC–DSGC connectivity found direction-selective GABA but non-direction-selective acetylcholine currents, suggesting that the release of the two transmitters and/or their receptors are spatially segregated. The low number of preferred-direction-oriented synapses and a fraction of null-direction-oriented synapses could provide symmetric synaptic acetylcholine signalling. Alternatively, the effect of acetylcholine could be largely paracrine, which is supported by the extra-synaptic localization of nicotinic receptors⁴⁴ and the existence of both nicotinic and muscarinic acetylcholine receptors remote from the choline-acetyltransferase-positive layers and thus distant to SAC dendrites^{34,45}, the sole source of acetylcholine in the retina⁶. The tight co-fasciculation of DSGC and SAC dendrites means short distances and therefore fast diffusion, in agreement with the fast rise times and latencies observed for cholinergic connections¹⁷.

The ultrastructure of synapses in the direction-selectivity circuit has been previously studied^{5,23}, but only the ability to collect a large enough EM volume at sufficient resolution made it possible to also study the directionality of presynaptic SAC dendrites. Combining, in the same piece of tissue, prior functional imaging and SBEM imaging allowed us to both target DSGCs for reconstruction and make a direct correlation between circuit structure and function. To trace the thin dendrites of SACs across hundreds of micrometres, the high z-resolution (23-nm section thickness) proved crucial. Furthermore, the prevalence of close encounters between neurites makes it unlikely (Fig. 3c–e) that a reliable determination of connectivity in the direction-selectivity circuit could ever be performed by fluorescent co-localization, but instead requires EM resolution.

The approach used here—tracing a cell and then going on to trace synaptically connected cells—is a virtual version of trans-synaptic-virus tracing^{46,47}. Although not (yet) providing the spatial reach of real virus methods, our approach, unlike those, does reveal the strength of connections in terms of the number of contacts, with an analysis of contact areas possible when needed. Whereas the current study relied

on targeted sparse reconstructions, similar SBEM data can, and probably will, be used for the dense (connectomic) reconstruction of the complete retinal circuitry.

METHODS SUMMARY

Calcium imaging. A dark-adapted retina from a wild-type C57BL/6 mouse was filled with the calcium indicator Oregon Green 488 BAPTA-1 by parallel-plate electroporation and imaged in a custom-built two-photon microscope during through-the-objective stimulation with bright bars moving in eight different directions at 0.5 mm s^{-1} . Sulphorhodamine 101 fluorescence was used to visualize the vasculature. Calcium imaging data were analysed using custom Matlab routines.

Tissue preparation for SBEM. The retina was incubated in 5 mg ml^{-1} horseradish peroxidase before fixation in 0.1 M cacodylate buffer with 4% sucrose and 2% glutaraldehyde, DAB precipitation in phosphate buffer with 4% sucrose, 5.6 mM DAB and 1.4 mM H_2O_2 , staining in 2% aqueous osmium tetroxide, enhancing by en bloc lead citrate, ethanol-series dehydration, and embedding in Epon hard. A second, not calcium-imaged, retina was more conventionally stained.

SBEM imaging and data analysis. The two-photon imaged region was mounted in a custom-built ultra-microtome operating inside the chamber of a field-emission scanning electron microscope (FEI QuantaFEG 200) and serial block-face imaged in 50 Pa water vapour at 2.5 keV landing energy at a dose of 10 electrons per nm^2 and a resolution of $16.5 \times 16.5 \times 23 \text{ nm}^3$ (for the conventionally stained sample, see Methods). A custom-designed back-scattered electron detector was used. SBEM data were aligned and stitched using custom Matlab routines. Skeletons were manually traced using custom written (KNOSSOS) software. Volumes were traced using ITK-SNAP.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 October 2010; accepted 10 January 2011.

- Barlow, H. B., Hill, R. M. & Levick, W. R. Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. *J. Physiol. (Lond.)* **173**, 377–407 (1964).
- Barlow, H. B. & Levick, W. R. The mechanism of directionally selective units in rabbit's retina. *J. Physiol. (Lond.)* **178**, 477–504 (1965).
- Taylor, W. R. & Vaney, D. I. Diverse synaptic mechanisms generate direction selectivity in the rabbit retina. *J. Neurosci.* **22**, 7712–7720 (2002).
- Fried, S. I., Munch, T. A. & Werblin, F. S. Mechanisms and circuitry underlying directional selectivity in the retina. *Nature* **420**, 411–414 (2002).
- Famiglietti, E. V. Synaptic organization of starburst amacrine cells in rabbit retina: analysis of serial thin sections by electron microscopy and graphic reconstruction. *J. Comp. Neurol.* **309**, 40–70 (1991).
- Tauchi, M. & Masland, R. H. The shape and arrangement of the cholinergic neurons in the rabbit retina. *Proc. R. Soc. Lond. B* **223**, 101–119 (1984).
- Yoshida, K. et al. A key role of starburst amacrine cells in originating retinal directional selectivity and optokinetic eye movement. *Neuron* **30**, 771–780 (2001).
- O'Malley, D. M., Sandell, J. H. & Masland, R. H. Co-release of acetylcholine and GABA by the starburst amacrine cells. *J. Neurosci.* **12**, 1394–1408 (1992).
- Euler, T., Detwiler, P. B. & Denk, W. Directionally selective calcium signals in dendrites of starburst amacrine cells. *Nature* **418**, 845–852 (2002).
- Chiao, C. C. & Masland, R. H. Starburst cells nondirectionally facilitate the responses of direction-selective retinal ganglion cells. *J. Neurosci.* **22**, 10509–10513 (2002).
- Grzywacz, N. M., Tootle, J. S. & Amthor, F. R. Is the input to a GABAergic or cholinergic synapse the sole asymmetry in rabbit's retinal directional selectivity? *Vis. Neurosci.* **14**, 39–54 (1997).
- Taylor, W. R. & Vaney, D. I. New directions in retinal research. *Trends Neurosci.* **26**, 379–385 (2003).
- Demb, J. B. Cellular mechanisms for direction selectivity in the retina. *Neuron* **55**, 179–186 (2007).
- Schachter, M. J. et al. Dendritic spikes amplify the synaptic signal to enhance detection of motion in a simulation of the direction-selective ganglion cell. *PLOS Comput. Biol.* **6** (2010).
- Borg-Graham, L. J. The computation of directional selectivity in the retina occurs presynaptic to the ganglion cell. *Nature Neurosci.* **4**, 176–183 (2001).
- Wei, W., Hamby, A. M., Zhou, K. & Feller, M. B. Development of asymmetric inhibition underlying direction selectivity in the retina. *Nature* **469**, 402–406 (2010).
- Lee, S., Kim, K. & Zhou, Z. J. Role of ACh-GABA co-transmission in detecting image motion and motion direction. *Neuron* **68**, 1159–1172 (2010).
- Yonehara, K. et al. Spatially asymmetric reorganization of inhibition establishes a motion-sensitive circuit. *Nature* **469**, 407–410 (2010).
- Mumm, J. S. et al. Laminar circuit formation in the vertebrate retina. *Prog. Brain Res.* **147**, 155–169 (2005).
- Famiglietti, E. V. A structural basis for omnidirectional connections between starburst amacrine cells and directionally selective ganglion cells in rabbit retina, with associated bipolar cells. *Vis. Neurosci.* **19**, 145–162 (2002).

21. Dong, W. *et al.* Dendritic relationship between starburst amacrine cells and direction-selective ganglion cells in the rabbit retina. *J. Physiol. (Lond.)* **556**, 11–17 (2004).
22. Chen, Y. C. & Chiao, C. C. Symmetric synaptic patterns between starburst amacrine cells and direction selective ganglion cells in the rabbit retina. *J. Comp. Neurol.* **508**, 175–183 (2008).
23. Dacheux, R. F., Chimento, M. F. & Amthor, F. R. Synaptic input to the on-off directionally selective ganglion cell in the rabbit retina. *J. Comp. Neurol.* **456**, 267–278 (2003).
24. White, J. G. *et al.* The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil. Trans. R. Soc. Lond.* **314**, 1–340 (1986).
25. Denk, W. & Horstmann, H. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biol.* **2**, e329 (2004).
26. Oyster, C. W., Amthor, F. R. & Takahashi, E. S. Dendritic architecture of ON-OFF direction-selective ganglion cells in the rabbit retina. *Vision Res.* **33**, 579–608 (1993).
27. Yang, G. & Masland, R. H. Receptive fields and dendritic structure of directionally selective retinal ganglion cells. *J. Neurosci.* **14**, 5267–5280 (1994).
28. Denk, W. & Detwiler, P. B. Optical recording of light-evoked calcium signals in the functionally intact retina. *Proc. Natl Acad. Sci. USA* **96**, 7035–7040 (1999).
29. Blankenship, A. G. *et al.* Synaptic and extrasynaptic factors governing glutamatergic retinal waves. *Neuron* **62**, 230–241 (2009).
30. Stosiek, C. *et al.* *In vivo* two-photon calcium imaging of neuronal networks. *Proc. Natl Acad. Sci. USA* **100**, 7319–7324 (2003).
31. Denk, W., Strickler, J. H. & Webb, W. W. Two-photon laser scanning fluorescence microscopy. *Science* **248**, 73–76 (1990).
32. Euler, T. *et al.* Eyecup scope—optical recordings of light stimulus-evoked fluorescence signals in the retina. *PLoS Arch.* **457**, 1393–1414 (2009).
33. Oyster, C. W. & Barlow, H. B. Direction-selective units in rabbit retina: distribution of preferred directions. *Science* **155**, 841–842 (1967).
34. Yamada, E. S. *et al.* Synaptic connections of starburst amacrine cells and localization of acetylcholine receptors in primate retinas. *J. Comp. Neurol.* **461**, 76–90 (2003).
35. Keeley, P. W. *et al.* Dendritic spread and functional coverage of starburst amacrine cells. *J. Comp. Neurol.* **505**, 539–546 (2007).
36. Hausselt, S. E. *et al.* A dendrite-autonomous mechanism for direction selectivity in retinal starburst amacrine cells. *PLoS Biol.* **5**, e185 (2007).
37. Lee, S. & Zhou, Z. J. The synaptic mechanism of direction selectivity in distal processes of starburst amacrine cells. *Neuron* **51**, 787–799 (2006).
38. Oesch, N. W. & Taylor, W. R. Tetrodotoxin-resistant sodium channels contribute to directional responses in starburst amacrine cells. *PLoS ONE* **5**, e12447 (2010).
39. He, S., Jin, Z. F. & Masland, R. H. The nondiscriminating zone of directionally selective retinal ganglion cells: comparison with dendritic structure and implications for mechanism. *J. Neurosci.* **19**, 8049–8056 (1999).
40. Kittila, C. A. & Massey, S. C. Effect of ON pathway blockade on directional selectivity in the rabbit retina. *J. Neurophysiol.* **73**, 703–712 (1995).
41. Caldwell, J. H., Daw, N. W. & Wyatt, H. J. Effects of picrotoxin and strychnine on rabbit retinal ganglion cells: lateral interactions for cells with more complex receptive fields. *J. Physiol. (Lond.)* **276**, 277–298 (1978).
42. Vaney, D. I. & Young, H. M. GABA-like immunoreactivity in cholinergic amacrine cells of the rabbit retina. *Brain Res.* **438**, 369–373 (1988).
43. Fried, S. I., Munch, T. A. & Werblin, F. S. Directional selectivity is formed at multiple levels by laterally offset inhibition in the rabbit retina. *Neuron* **46**, 117–127 (2005).
44. Dmitrieva, N. A. *et al.* Identification of cholinergic glycinergic neurons in the mammalian retina. *J. Comp. Neurol.* **456**, 167–175 (2003).
45. Dmitrieva, N. A., Strang, C. E. & Keyser, K. T. Expression of $\alpha 7$ nicotinic acetylcholine receptors by bipolar, amacrine, and ganglion cells of the rabbit retina. *J. Histochem. Cytochem.* **55**, 461–476 (2007).
46. Wickersham, I. R. *et al.* Retrograde neuronal tracing with a deletion-mutant rabies virus. *Nature Methods* **4**, 47–49 (2007).
47. Granstedt, A. E. *et al.* Fluorescence-based monitoring of *in vivo* neural activity using a circuit-tracing pseudorabies virus. *PLoS ONE* **4**, e6923 (2009).
48. Briggman, K. L. & Euler, T. Bulk electroporation and population calcium imaging in the adult mammalian retina. *J. Neurophysiol.* (in the press).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Euler for many useful discussions and help with the functional imaging experiments. We also thank H. Horstmann and S. Mikula for help with staining procedures; J. Kornfeld and F. Svava for programming KNOSSOS; J. Tritthardt for developing electronic circuits and M. Müller for help with the acquisition software; J. Hanne, H. Jakobi and H. Wissler for help with training tracers; M. Feller and Z. J. Zhou for discussion of their results; and J. Bollmann, A. Karpova and S. Seung for comments on the manuscript. We thank N. Abazova, E. Abs, A. Antunes, P. Bastians, M. Beining, J. Buhmann, F. Drawitsch, L. Ehm, F. Isensee, H. Jakobi, S. Kaspar, A. Khan, M. Kiapes, A. Klein, S. Laiouar, E. Möller, J. Trendel, P. Weber, K. Weiß, E. Wiegand and H. Wissler for the tracing work.

Author Contributions K.L.B. and W.D. designed the study; W.D. designed the microtome; K.L.B. performed the calcium imaging and SBEM experiments, K.L.B. and M.H. analysed data; K.L.B., M.H. and W.D. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.L.B. (briggman@mpimf-heidelberg.mpg.de).

METHODS

Animals and tissue preparation. Two adult wild-type (C57BL/6) mice were used, one for the calcium-imaging experiment (P29), and one prepared for conventional staining (P30). The mouse used for calcium imaging had been dark-adapted for 2 h before the experiment, with all subsequent procedures up to the functional imaging carried out under dim red illumination. The mice were anaesthetized with Isoflurane (Baxter) inhalation and killed by cervical dislocation. The eyes were enucleated and transferred to a dish containing carboxygenated room-temperature saline, in which the retinas were dissected. All procedures were approved by the local animal care committee and were in accordance with the law of animal experimentation issued by the German Federal Government.

Physiological solutions and calcium-indicator loading. As standard medium we used for tissue dissection, electroporation and recordings a commercially available saline (Biometra) that was supplemented with 0.5 mM L-glutamine and carboxygenated (95% O₂/5% CO₂). To visualize the retinal morphology and the vasculature under the two-photon microscope^{32,49} we added sulphorhodamine 101 (SR101, final concentration 0.5–1 μ M; Sigma-Aldrich) to the superfusion medium. For the electroporation, a 5 mM solution (in our standard medium) of the synthetic calcium indicator dye Oregon Green 488 BAPTA-1 (OGB-1, hexapotassium salt, Invitrogen) was used. We loaded the ganglion cell layer of the retina by bulk electroporation as described elsewhere (K. L. Briggman and T. Euler, manuscript submitted). Briefly, half of one retina was whole-mounted on filter paper, photoreceptor-side down, and placed between two horizontal parallel-plate electrodes with a 5 μ l drop of OGB-1 covering the ganglion cell layer. The electroporation parameters were +13V/2 mm (applied voltage to upper electrode/distance between electrodes), 10-ms pulse width, 1-Hz pulse frequency, 10 square-wave pulses. The retina was immediately transferred to the heated (30–32 °C) recording chamber and allowed to recover for 45 min before two-photon imaging.

Two-photon microscopy. For two-photon imaging³¹ we used a custom-built microscope ('eyecup scope'³²), equipped with through-the-objective light stimulation and two detection channels for fluorescence imaging (red, HQ 622 BP 36, and green, D 535 BP 50 or 520 BP 30; AHF/Chroma). The excitation source was a mode-locked Ti:sapphire laser (Mira-900, Coherent) tuned to \approx 930 nm. The microscope was used to simultaneously visualize the retinal structure with SR-101 (red channel, see above) and to monitor calcium activity reflected by OGB-1 fluorescence changes (green channel). During functional imaging, the scan parameters were 64 \times 64 pixels at 7.8 Hz frame rate. A marker signal coming from the light stimulator (see next section) was embedded into the image data, at the beginning of each scan line, to allow the registration of calcium signals and stimulus presentations. The imaged field of view for each stimulus presentation was 100 μ m \times 100 μ m. We acquired data from nine such fields, arranged in a 3 \times 3 mosaic, for a total recorded area of 300 μ m \times 300 μ m.

Light stimulation. Our light stimulator used custom-written software running under Windows XP to present spatially and temporally structured stimuli on an 800 \times 600 pixel miniature LCoS display (i-glasses, EST), alternately illuminated by two band-pass-filtered (blue, 400 BP 20; green, 578 BP 10; AHF/Chroma) LEDs within each frame, for details see ref. 32. We used 'grey' stimuli (with both blue and green LEDs at the same intensity) with a positive contrast, that is, bright foreground on darker background and \sim 200 and \sim 60 \times 10³ photons s⁻¹ μ m⁻², respectively. The LCoS display was coupled into the main optical path of the microscope and, hence, the stimuli were projected onto the retina through the objective lens (XLUMPlanFL 20 \times 0.95 NA water-immersion, Olympus). Each of the eight stimuli used was a bar (200 μ m wide \times 1,000 μ m long) moving in one of eight different directions at 0.5 mm s⁻¹.

Calcium-imaging data analysis. Images were analysed offline using custom Matlab (The Mathworks, Inc.) routines. In brief, the pixel intensities within circular regions of interest (ROIs) that were manually placed over each cell body were averaged for each frame. Fractional fluorescence changes were calculated using the mean intensity of each ROI across an entire recording trial as the baseline. The responses to individual stimulus presentations were extracted using the embedded marker signal from the stimulator. The responses were averaged across five stimulus presentations for each of the eight directions. Tuning curves were generated by integrating the total area under the responses and normalizing to the peak integrated area for each cell.

Sample preparation for electron microscopy. We prepared two retinas for electron microscopy; for the first, calcium-imaged retina (e2198) we used a surface-labelling protocol, for the second retina (k0563) we used a conventional staining protocol. Following calcium imaging, e2198 was removed from the recording chamber and placed in 2 ml of room-temperature saline containing 5 mg ml⁻¹ horseradish peroxidase (HRP; activity = 1,700 U mg⁻¹, Serva). After 10 min the HRP solution was removed and replaced with fixative solution containing 0.1 M cacodylate buffer, 4% sucrose and 2% glutaraldehyde, pH 7.2 (Serva). The tissue was fixed for 2 h at room temperature and then rinsed in

0.1 M cacodylate buffer + 4% sucrose overnight. A 1 \times 1 mm² region of the retina centred on the calcium-imaged region was excised and transferred to a solution containing 0.1 M phosphate buffer, 4% sucrose, 5.6 mM diaminobenzidine (DAB; Serva), and 1.4 mM hydrogen peroxide and incubated for 8 h at room temperature. After a 2 h rinse in 0.1 M cacodylate buffer + 4% sucrose, the sample was stained in 2% aqueous osmium tetroxide (Serva) for 2 h at room temperature, and then rinsed in distilled water for 2 h. The osmium-labelled DAB product was further enhanced by en-bloc lead citrate staining, prepared as in ref. 50. The tissue was dehydrated through an ethanol series (70%, 90%, 100%), transferred to propylene oxide, infiltrated with 50%/50% propylene oxide/Epon Hard⁵¹, and then 100% Epon Hard. The block was cured at 60 °C for 24 h.

For conventional staining (k0563), the retina was fixed in 0.1 M cacodylate buffer, 4% sucrose and 2% glutaraldehyde, pH 7.2 for 2 h, then rinsed overnight in 0.1 M cacodylate buffer and 4% sucrose. The tissue was then stained in a solution containing 1% osmium tetroxide, 1.5% potassium ferrocyanide, and 0.15 M cacodylate buffer for 2 h at room temperature⁵². The osmium stain was amplified⁵³ with 1% thiocarbonylhydrazide (1 h at 50 °C), and 2% osmium tetroxide (1 h at room temperature). The tissue was then stained with 2% aqueous uranyl acetate for 12 h at room temperature and lead aspartate⁵⁴ for 12 h at room temperature. The dehydration and embedding procedures were as described above.

SBEM acquisition. The retinas were cut out of the flat-embedding blocks and re-embedded in Epon Hard, on aluminium stubs for SBEM, with the retinal plane vertical. The samples were then trimmed to a block face of \sim 200 μ m wide and \sim 400 μ m long. Care was taken to trim down the surface of e2198 to just outside the calcium-imaged region, using vasculature landmarks. The samples were imaged in a scanning electron microscope with a field-emission cathode (QuantaFEG 200, FEI Company). Back-scattered electrons were detected using a custom-designed detector based on a special silicon diode (AXUV, International Radiation Detectors) combined with a custom-built current amplifier. The incident electron beam had an energy of 2.5 keV and a current of \sim 230 pA for e2198 and an energy of 2.0 keV and a current of \sim 100 pA for k0563. At a pixel dwell time of 1.9 μ s and a pixel size of 16.5 nm \times 16.5 nm (e2198) and 5 μ s and 12 nm \times 12 nm (k0563), this corresponds to doses of about 10 (e2198) and 22 (k0563) electrons per nm², not accounting for skirting due to low vacuum operation. To prevent charging for e2198, the chamber was kept at a pressure of 50 Pa of water vapour; k0563, which had a higher intrinsic conductivity, was imaged at high vacuum, with the sides of the block evaporation-coated with a 100–200-nm-thick layer of gold. The electron microscope was equipped with a custom-made microtome that is similar to the one described in ref. 25 and is the basis of a commercially available instrument (3view, Gatan). Our microtome allows the repeated removal of the block surface at increments as small as \sim 23 nm, which was the cutting thickness for e2198 (\sim 25 nm for k0563). 13,000 (for e2198) and 5,760 (for k0563) consecutive block faces were imaged, resulting in data volumes of 3,584 \times 21,658 \times 13,000 voxels (1 \times 7 mosaic of 3,584 \times 3,094 images, e2198) and 4,096 \times 5,304 \times 5,760 voxels (2 \times 3 mosaic of 2,048 \times 1,768 images, k0563), corresponding to approximate spatial volumes of 60 \times 350 \times 300 μ m³ and 50 \times 65 \times 145 μ m³. The edges of neighbouring mosaic images overlapped by \sim 1 μ m. For both retinas, the imaged region spanned the inner plexiform layer of the retina and included the ganglion cell layer and part of the inner nuclear layer. Cross-correlation-derived shift vectors between neighbouring mosaic images and consecutive slices were used for a global least-squares fit across all shift vectors to align the data sets off-line to subpixel precision by Fourier shift-based interpolation. The data sets were then split into cubes (128 \times 128 \times 128 voxels) for viewing in KNOSSOS.

Skeleton tracing and contact annotation. Skeletons were traced using KNOSSOS (details in M. Helmstaedter, K. L. Briggman and W. Denk, manuscript submitted), and consisted of nodes and connections between them. Nodes could be annotated as needed and were placed approximately every 250 nm. Tracers were instructed to pay attention to branch points and to value accuracy over speed. Typically the cross-sectional view closest to being orthogonal to its axis was used to follow the neurite. The DSGCs were traced by an expert tracer and SACs were traced by trained tracers. Tracers of the very thin SAC dendrites were instructed to skeletonize only those dendrites that could be unequivocally followed. Some branches were, therefore, missed at difficult locations in the data set. As a consequence, some of the SAC skeletons are incomplete (Supplementary Fig. 7). To find potential contact sites, locations were identified where skeletons approached each other more closely than 1.5 μ m. Those locations where then manually inspected and classified into putative synapses or incidental contacts. Some of the local contacts were manually volume traced using ITK-SNAP⁵⁵.

49. Schlichtenbrede, F. C. *et al.* Toxicity assessment of intravitreal triamcinolone and bevacizumab in a retinal explant mouse model using two-photon microscopy. *Invest. Ophthalmol. Vis. Sci.* **50**, 5880–5887 (2009).

50. Fahmy, A. An Extemporaneous Lead Citrate Stain for Electron Microscopy 148–149 (Proc. 25th Annu. EMSA Meeting, 1967).

51. Glauert, A. M. & Lewis, P. R. Biological specimen preparation for transmission electron microscopy. In *Practical Methods in Electron Microscopy* xxi (Princeton Univ. Press, 1998).
52. Karnovsky, M. J. *Use of Ferrocyanide-reduced osmium in electron microscopy* 146 (Proc. 14th Annual Meeting Am. Soc. Cell Biol., 1971).
53. Seligman, A. M., Wasserkrug, H. L. & Hanker, J. S. A new staining method (OTO) for enhancing contrast of lipid-containing membranes and droplets in osmium tetroxide-fixed tissue with osmiophilic thiocarbohydrazide (TCH). *J. Cell Biol.* **30**, 424–432 (1966).
54. Walton, J. Lead aspartate, an en bloc contrast stain particularly useful for ultrastructural enzymology. *J. Histochem. Cytochem.* **27**, 1337–1342 (1979).
55. Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* **31**, 1116–1128 (2006).

Inactivating mutations of acetyltransferase genes in B-cell lymphoma

Laura Pasqualucci^{1,2}, David Dominguez-Sola¹, Annalisa Chiarenza¹, Giulia Fabbri¹, Adina Grunn¹, Vladimir Trifonov³, Lawryn H. Kasper⁴, Stephanie Lerach⁴, Hongyan Tang¹, Jing Ma⁵, Davide Rossi⁶, Amy Chadburn⁷, Vundavalli V. Murty^{1,2}, Charles G. Mullighan⁸, Gianluca Gaidano⁶, Raul Rabadan³, Paul K. Brindle⁴ & Riccardo Dalla-Favera^{1,2,9}

B-cell non-Hodgkin's lymphoma comprises biologically and clinically distinct diseases the pathogenesis of which is associated with genetic lesions affecting oncogenes and tumour-suppressor genes. We report here that the two most common types—follicular lymphoma and diffuse large B-cell lymphoma—harbour frequent structural alterations inactivating *CREBBP* and, more rarely, *EP300*, two highly related histone and non-histone acetyltransferases (HATs) that act as transcriptional co-activators in multiple signalling pathways. Overall, about 39% of diffuse large B-cell lymphoma and 41% of follicular lymphoma cases display genomic deletions and/or somatic mutations that remove or inactivate the HAT coding domain of these two genes. These lesions usually affect one allele, suggesting that reduction in HAT dosage is important for lymphomagenesis. We demonstrate specific defects in acetylation-mediated inactivation of the BCL6 oncoprotein and activation of the p53 tumour suppressor. These results identify *CREBBP/EP300* mutations as a major pathogenetic mechanism shared by common forms of B-cell non-Hodgkin's lymphoma, with direct implications for the use of drugs targeting acetylation/deacetylation mechanisms.

Diffuse large B-cell lymphoma (DLBCL) represents the most common form of B-cell non-Hodgkin lymphoma (B-NHL), accounting for about 30% of new diagnoses and also arising as a frequent clinical evolution of follicular lymphoma¹. The molecular pathogenesis of DLBCL is associated with multiple genetic lesions that segregate in part with individual phenotypic subtypes, namely germinal-centre B-cell-like (GCB) and activated B-cell-like (ABC) DLBCL, suggesting the involvement of distinct oncogenic pathways^{2–8}. However, the full spectrum of lesions that contribute to malignant transformation remains unknown. Genome-wide efforts toward the identification and functional characterization of the entire set of structural alterations present in the DLBCL genome are required for a complete understanding of its pathogenesis⁹.

Towards this end, we have integrated next-generation whole-exome sequencing analysis of seven DLBCL cases and genome-wide high-density single nucleotide polymorphism (SNP) array analysis of 72 DLBCL cases. This combined approach led to the identification of over 450 loci that are affected by somatic point mutations and/or by recurrent, focal gene copy number aberrations. Among those that have been independently validated, the most commonly involved regions were those harbouring the acetyltransferase genes *CREBBP* (*CBP*) and *EP300* (*p300*). *CREBBP* encodes a highly conserved and ubiquitously expressed nuclear phosphoprotein that, together with the closely related protein EP300, belongs to the KAT3 family of histone/protein lysine acetyltransferases^{10,11}. *CREBBP* and EP300 function as transcriptional coactivators for a large number of DNA-binding transcription factors involved in multiple signalling and

developmental pathways, by modifying lysine residues on both histone and non-histone nuclear proteins^{12,13}.

CREBBP and EP300 enhance transcription by multiple mechanisms, including: (1) targeted acetylation of chromatin^{12,13}; (2) acetylation of transcriptional activators (such as the tumour suppressors p53 and GATA1)^{14–17}; and (3) acetylation-mediated inactivation of transcriptional repressors (such as the DLBCL-associated oncogene BCL6)¹⁸. Additionally, both molecules were found to exhibit ubiquitin ligase activity^{19,20}. Consistent with the involvement in critical cellular functions, homozygous null mice for either *Crebbp* or *Ep300* are early embryonic lethal^{21,22}, and the same is true for compound *Crebbp/Ep300* double heterozygous mice^{21,22}. Haploinsufficiency of *CREBBP* (and, rarely, *EP300*) is responsible for Rubinstein–Taybi syndrome, an autosomal congenital disorder characterized by mental and growth retardation, skeletal abnormalities and predisposition to tumour development^{23–25}. Somatic mutations of these two genes are exceedingly rare in epithelial cancers^{26–28}, and only three cases were reported to carry *EP300*, but not *CREBBP* mutations in haematologic malignancies^{29,30}. Additionally, chromosomal translocations of *CREBBP* are associated with an infrequent type of acute myeloid leukaemia and with therapy-related acute myeloid leukaemia and myelodysplastic syndrome, although the precise consequences of these lesions have not been fully elucidated^{31–33}.

Monoallelic lesions of *CREBBP* in DLBCL

Following initial observations from whole-exome sequencing analysis of seven DLBCL cases and paired normal DNAs, we performed targeted

¹Institute for Cancer Genetics and the Herbert Irving Comprehensive Cancer Center, Columbia University, New York, New York 10032, USA. ²Department of Pathology and Cell Biology, Columbia University, New York, New York 10032, USA. ³Department of Biomedical Informatics and the Center for Computational Biology and Bioinformatics, Columbia University, New York, New York 10032, USA. ⁴Department of Biochemistry, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ⁵The Hartwell Center for Bioinformatics and Biotechnology, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ⁶Division of Hematology, Department of Clinical and Experimental Medicine and IRCAD, Amedeo Avogadro University of Eastern Piedmont, Novara, 28100, Italy. ⁷Department of Pathology, Northwestern University, Feinberg School of Medicine, Chicago, Illinois 60611, USA. ⁸Department of Pathology, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ⁹Department of Genetics and Development, Columbia University, New York, New York 10032, USA.

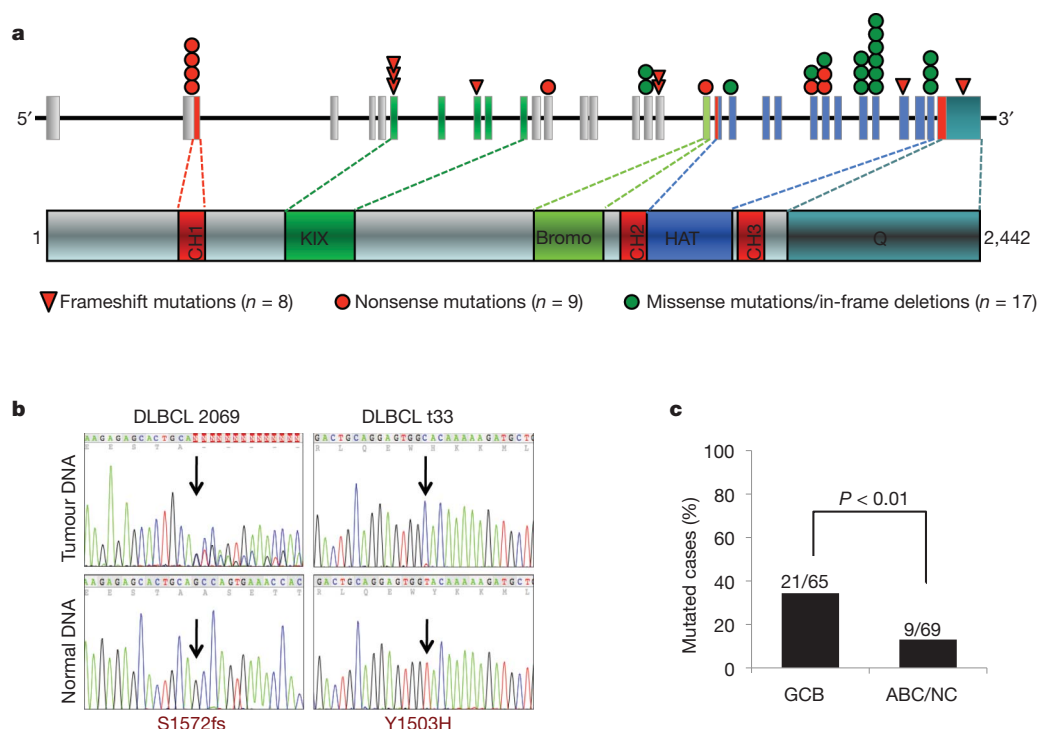


Figure 1 | The *CREBBP* gene is mutated in DLBCL. **a**, Schematic diagram of the *CREBBP* gene (top) and protein (bottom). Exons are colour-coded according to the corresponding protein functional domains (CH, cysteine-histidine rich; KIX, CREB-binding; Bromo, bromodomain; HAT, histone acetyltransferase; Q, polyglutamine stretch). Colour-coded symbols depict

distinct types of mutations. **b**, Sequencing traces of representative mutated DLBCL tumour samples and paired normal DNA; arrows point to the position of the nucleotide change (amino acid change shown at the bottom). **c**, Distribution of *CREBBP* mutations in major DLBCL subtypes; the actual number of mutated samples over the total analysed is given. NC, not classified.

re-sequencing of the entire *CREBBP* coding exons in 134 DLBCL samples representative of major phenotypic subtypes. This analysis revealed a total of 34 sequence variants distributed over 30 samples and the somatic origin of which was documented by analysis of paired normal DNAs (available in eight cases) (Fig. 1a and b, and Supplementary Table 1). Of these variants, 17 (50%) were inactivating events, including nonsense mutations ($n = 9$), frameshift insertions/deletions ($n = 7$) and mutations at consensus splice donor/splice acceptor sites ($n = 1$), which generate aberrant transcripts carrying premature stop codons. On the basis of their distribution along the *CREBBP* protein, these mutations are all predicted to cause the elimination or truncation of the HAT domain (Fig. 1a and Supplementary Table 1). The remaining variants included three in-frame deletions and 14 missense mutations, primarily within the HAT domain (Fig. 1a and Supplementary Table 1), suggesting that they may be functionally important (see below). Although *CREBBP* mutations were identified in both DLBCL phenotypic subtypes, their frequency was significantly higher in GCB-DLBCL, where they account for about 32% of the cases ($n = 21/65$) as compared to 13% in ABC/non-classified (NC) DLBCL ($n = 9/69$; $P < 0.01$) (Fig. 1c).

High-density SNP array analysis, available for 72 samples from the same panel, and fluorescence *in situ* hybridization (FISH) analysis revealed the presence of monoallelic deletions encompassing or internal to the *CREBBP* locus in eight additional cases (five GCB-DLBCL and three ABC/non-GC-DLBCL), and a homozygous deletion in one patient (Fig. 2a and b, and Supplementary Fig. 1). Notably, the loss of genetic material was smaller than 240 kilobases (kb) in four cases and, in two patients (2147 and 2043), involved only a limited subset of *CREBBP* exons, thereby identifying this gene as the specific target (Fig. 2a and Supplementary Table 2). In two additional samples for which copy number data were not available, direct sequencing analysis revealed a hemizygous missense mutation, reflecting either the loss of the second allele or copy neutral loss of heterozygosity

(Supplementary Table 1). When combining the sequencing data with the copy number data, *CREBBP* mutations and deletions were found to be mutually exclusive in most samples, revealing a predominantly monoallelic distribution ($n = 33/39$ cases) (Fig. 2c). Rare instances of biallelic genetic lesions include the homozygous loss, a missense mutation with loss of heterozygosity (two primary biopsies), biallelic nucleotide substitutions (two cases), and a frameshift deletion with missense mutation of the second allele in the OCI-Ly8 cell line. Thus, 29% of all DLBCL patients ($n = 39/134$), corresponding to 41.5% of GCB-DLBCL and 17% of ABC-DLBCL, harbour genomic alterations affecting the *CREBBP* gene (Fig. 2d).

Frequent mutations of *CREBBP* in follicular lymphoma

We next analysed various types of mature B-NHL, including follicular lymphoma, Burkitt lymphoma, marginal-zone lymphoma and chronic lymphocytic leukaemia. Mutations analogous to those found in DLBCL were frequent in follicular lymphoma (32.6%, with 16 events distributed over 15/46 cases), but not in other lymphoma types, suggesting a specific role in the pathogenesis of these two diseases (Supplementary Fig. 2). This analysis also revealed the existence of several mutational hotspots at specific codons within the HAT domain, including R1446 (also mutated in B-cell acute lymphoblastic leukaemia³⁴), Y1503 and D1435; in addition, a 3-base-pair (bp) in-frame deletion causing the loss of S1680 was observed in five cases, suggesting a functional role for this currently uncharacterized serine (Supplementary Table 1). None of these recurrent changes was detected in paired normal DNA, indicating that they do not represent germline polymorphisms. Although copy number data were not available for the same follicular lymphoma panel, array-CGH analysis performed on a distinct data set showed deletions spanning the *CREBBP* locus in only 1/68 cases (not shown). Collectively, these findings identify somatic mutations of *CREBBP* as a common event in follicular lymphoma.

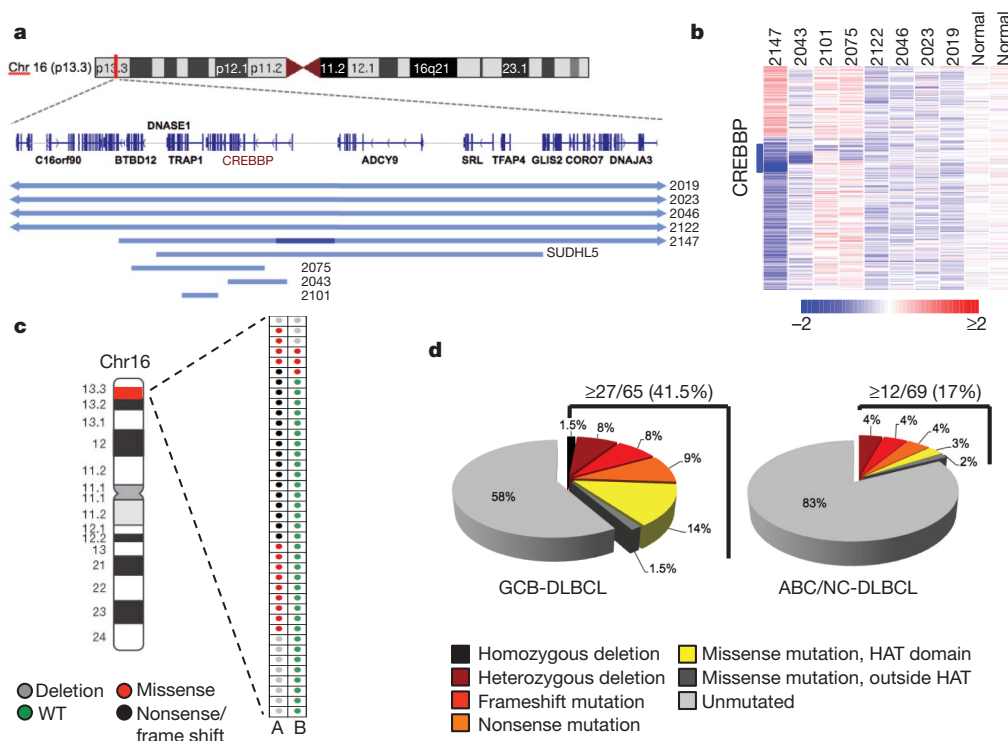


Figure 2 | Mutations and deletions of *CREBBP* are predominantly monoallelic. **a**, Map of the genomic region encompassing *CREBBP* and its neighbouring genes on chromosome 16p13.3. Blue lines below the map indicate the extent of the deletions identified in nine DLBCL samples, with the darker blue segment corresponding to a homozygous loss. **b**, dChipSNP heatmap showing median-smoothed log₂ copy number ratio for eight DLBCL biopsies harbouring *CREBBP* deletions, and two normal DNAs. A vertical blue bar

indicates the location of the *CREBBP* locus; in the red-to-blue colour scale, white corresponds to a normal (diploid) copy number log-ratio, blue corresponds to deletion and red to gain. **c**, Allelic (A or B) distribution of *CREBBP* genetic lesions in individual DLBCL samples. **d**, Overall frequency of *CREBBP* structural alterations in DLBCL subtypes (mutations and deletions, combined).

EP300 mutations in DLBCL and follicular lymphoma

Given the significant structural and functional similarities between *CREBBP* and *EP300*, we investigated whether this second member of the KAT3 acetyltransferase family is also targeted by structural alterations in B-NHL. Mutational analysis of the same panel identified 19 sequence variants leading to amino acid changes ($n = 11$), in-frame deletions ($n = 2$), and premature stop codons due to frameshift deletions, aberrant splicing or nonsense mutations ($n = 6$) (Supplementary Fig. 3a and Supplementary Table 3). These lesions were found in 10% of DLBCL ($n = 13/134$) and 8.7% of follicular lymphoma samples ($n = 4/46$), but were almost absent in other B-NHLs (Supplementary Fig. 3b). Seven additional DLBCLs harboured monoallelic deletions spanning, although not limited to, the *EP300* locus (Supplementary Fig. 3c, Supplementary Table 4 and data not shown). Notably, structural alterations of *CREBBP* and *EP300* coexist in only a minority of the affected cases ($n = 6/53$ DLBCL and $0/19$ follicular lymphoma) (Supplementary Fig. 3d), suggesting that inactivation of these loci is at least in part functionally equivalent. Thus, overall about 39% of all DLBCL and at least 41% of follicular lymphoma cases (based on mutations only) display structural alterations of KAT3 family genes.

CREBBP and EP300 protein expression

To compare the *CREBBP* and *EP300* protein levels in normal and transformed B cells, and to investigate the expression of the retained normal allele in cases carrying monoallelic genomic alterations, we examined the expression pattern of these two proteins in reactive human tonsils and in 78 DLBCL primary cases, of which 49 harboured both genes in wild-type configuration, using immunofluorescence and immunohistochemical analysis. Consistent with their reported ubiquitous expression, *CREBBP* and *EP300* were detectable in naive B cells

within the mantle zone and, at higher levels, in germinal-centre B cells (Fig. 3a). Notably, most of the monoallelically deleted DLBCL cases were positive for the two proteins, although at reduced levels, demonstrating that the residual wild-type allele is expressed (see Fig. 3b and c for representative examples). Furthermore, reverse-transcription polymerase chain reaction (RT-PCR) amplification and direct sequencing of the ten *CREBBP*-mutated cell lines invariably showed the presence of the wild-type allele (Supplementary Fig. 4), while western blot analysis using antibodies directed against the amino-terminal portion of *CREBBP* revealed the expression of a full-length protein, corresponding to the wild-type allele, in most of the cell lines carrying truncating gene mutations, as well as in the monoallelically deleted *SUDHL5* cells (Fig. 3d, top panel). Slightly different findings were observed for *EP300*, where the presence of truncating mutations or deletions was associated with the absence of protein expression in four affected cell lines (Fig. 3d, second panel from top). With the exception of *SUDHL2*, both *CREBBP* and *EP300* messenger RNAs were readily detected by northern blot analysis in all lines analysed (Fig. 3d, panels 4 and 5 from top).

Interestingly, a few additional lines were found to express very low or undetectable protein levels, despite the presence of mRNA and the absence of structural alterations in the corresponding gene (see *CREBBP* in BJAB and *SUDHL2*, or *EP300* in *SUDHL7*) (Fig. 3d). Similarly, 6/78 (8%) DLBCL biopsies from patients with intact alleles appeared to lack expression of the two proteins, either simultaneously (three cases) or individually (two *CREBBP*⁺/*EP300*[−] cases and one *CREBBP*[−]/*EP300*⁺ case) (Fig. 3c and e). Thus, the fraction of DLBCL patients with defective *CREBBP* and *EP300* may be higher than that determined on the basis of genetic lesions alone, suggesting that alternative mechanisms of KAT3 gene family inactivation may have a role in this disease.

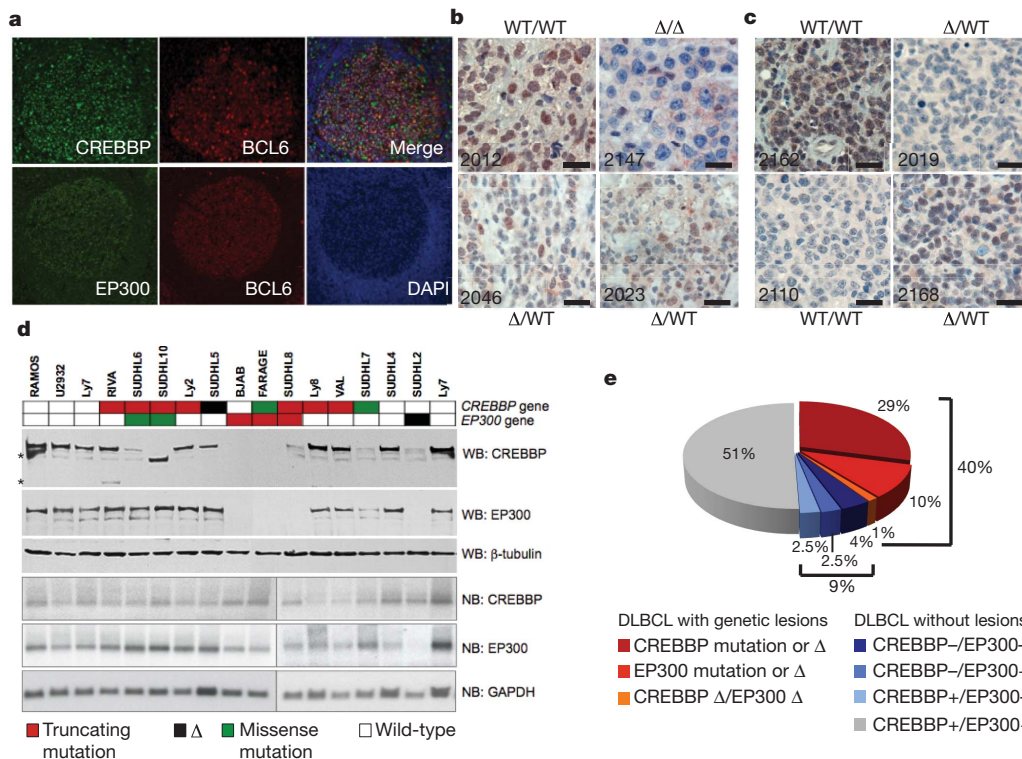


Figure 3 | CREBBP and EP300 expression in normal and transformed B-cells. **a**, Immunofluorescence analysis of reactive tonsils. BCL6 identifies germinal-centre B cells, and DAPI is used to detect nuclei. **b**, **c**, Immunohistochemistry analysis of CREBBP (**b**) and EP300 (**c**) protein expression in representative DLBCL biopsies (genomic status as indicated, Δ is deletion; scale bar, 100 μ m). Sample 2147, which harbours a homozygous *CREBBP* deletion, serves as negative control. **d**, Western blot (WB) and

northern blot (NB) analysis of DLBCL cell lines carrying wild-type or aberrant *CREBBP* and *EP300* alleles (colour-coded as indicated). The aberrant band in SUDHL10 corresponds in size to the predicted ~220-kilodalton (kDa) *CREBBP* truncated protein. *Non-specific bands. Tubulin and GAPDH control for total protein and RNA loading, respectively. **e**, Overall proportion of DLBCL biopsies showing defective *CREBBP*/*EP300* function due to genetic lesions (red scale) and/or lack of protein expression (blue scale).

CREBBP missense mutants fail to acetylate BCL6 and p53

The presence of gross gene deletions and the distribution of truncating mutations clearly predict a complete gene inactivation or the loss of multiple key functional domains, including the HAT; on the contrary, the functional consequences of the numerous *CREBBP* missense mutations required direct experimental analysis. Notably, of the 30 total events identified, 27 cluster within HAT coding exons, suggesting a selective pressure to alter the *CREBBP* enzymatic activity (Supplementary Table 1). In particular, 19 mutations were located within a 68-amino-acid stretch that is 96% identical to EP300 and includes the contact surface for coenzyme A (CoA)³⁵ (Supplementary Fig. 5). We therefore examined the effect of these mutations on the ability of *CREBBP* to acetylate known physiologic substrates. We selected BCL6 and p53 because of their biological relevance for germinal-centre development and lymphomagenesis^{4,36}. In fact, acetylation of the tumour suppressor p53 is indispensable for its transcriptional activity^{14,15,37}, whereas EP300-mediated acetylation of the proto-oncoprotein BCL6 leads to inactivation of its transcriptional repressor function¹⁸.

Transient transfection/co-immunoprecipitation assays confirmed that, analogous to EP300¹⁸, *CREBBP* binds to and acetylates BCL6, leading to a dose-dependent impairment in its ability to repress a BCL6-responsive reporter gene (Supplementary Fig. 6a and b). We then generated haemagglutinin (HA)-tagged constructs for expression of nine representative *CREBBP* alleles harbouring missense mutations within ($n = 6$) or immediately outside ($n = 2$) the core HAT domain, as well as a premature stop codon (R1360X) (Fig. 4a). Notably, all of the HAT domain mutant proteins had lost

their ability to acetylate BCL6 (Fig. 4b) and to interfere with its trans-repression activity, consistent with the reported role of acetylation in inactivating BCL6 (Fig. 4c). Conversely, no significant effects were observed from the C1240R and K1320R proteins or from two additional alleles (P1053L, Q1079H) harbouring mutations outside the HAT domain (Supplementary Table 1 and data not shown), although C1240R appeared to have lost its activity in the reporter assay, suggesting that alternative mechanisms may be involved.

When tested on the tumour suppressor p53, the same core HAT mutants were either impaired or severely attenuated (H1487Y and D1435E) in their acetylation activity, as compared to wild-type *CREBBP* or to C1240R and K1320R (Fig. 4d). It should be noted that C1240R and K1320R represent a second mutational event in cases carrying an additional truncating *CREBBP* mutation (see Supplementary Table 1), suggesting the possibility that they represent passenger events or that they confer more subtle functional consequences. Taken together, these findings demonstrate that *CREBBP* missense mutant alleles have been selected for their defective HAT activity. In particular, the impairment on BCL6 and p53 acetylation is consistent with specific effects of *CREBBP* genetic lesions in favouring the constitutive activity of the BCL6 oncogene over the function of the p53 tumour suppressor.

CREBBP mutations reduce affinity for acetyl-CoA binding

To investigate the mechanisms responsible for the observed loss of function on BCL6 and p53, we examined the effect of *CREBBP* mutations on subcellular localization, enzyme-substrate complex formation, and enzymatic activity. All mutants tested were correctly localized in

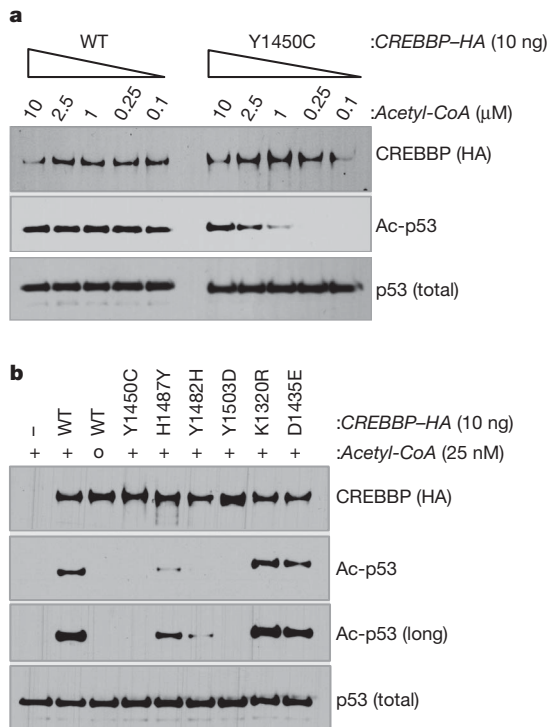


Figure 5 | DLBCL-associated mutations in the CREBBP HAT domain decrease its affinity for acetyl-CoA. **a**, Western blot analysis of *in vitro* acetyltransferase reactions performed using the WT or mutant Y1450C CREBBP-HA recombinant protein and a GST-p53 substrate in the presence of decreasing amounts of acetyl-CoA. Anti-HA and anti-p53 antibodies document the presence of equivalent amounts of effector and substrate proteins in the reaction. Ac, acetylated. **b**, *In vitro* acetyltransferase activity of the indicated CREBBP-HA mutant proteins in the same assay using 25 nM acetyl-CoA.

of these common diseases. Moreover, the observation of significantly reduced CREBBP and EP300 expression levels in a sizeable fraction of DLBCL samples, independent of genetic lesions (Fig. 3e and Supplementary Fig. 12), suggests that additional epigenetic mechanisms may cause reduction of HAT dosage in a larger fraction of tumours.

One key observation of this study is that *CREBBP/EP300* lesions are mostly detected in heterozygosity, suggesting a haploinsufficient role in tumour suppression. This notion is supported by several observations. First, congenital heterozygous mutations of *CREBBP/EP300* are sufficient to cause significant pathologic and developmental phenotypes, including tumorigenesis, in patients with Rubinstein-Taybi syndrome^{23–25}, thereby confirming the deleterious effect of reduced HAT activity. Second, a fraction of mice with conditional deletion of *Crebbp* in mature B cells showed reduced survival past 12 months of age⁴¹. Haematologic malignancies were also reported in 22% of constitutive 10–21-month-old *Crebbp* heterozygous mice and about 30% of chimeric animals after bone marrow or spleen cell transplantation from *Crebbp* heterozygotes⁴². Third, in contrast with the abundantly expressed histone deacetylases, HATs are limiting in the cell, suggesting that small dosage variations can have severe biological consequences⁴³. Overall, these data provide direct evidence that CREBBP (and, more rarely, EP300) are haploinsufficient tumour suppressors, the specific roles of which in lymphomagenesis will have to be tested by conditional deletion of these alleles in germinal-centre B cells.

Given the global involvement of CREBBP/EP300 on gene transcriptional regulation, it is difficult to predict which cellular targets/pathways may be critically affected by HAT reduction in lymphomagenesis. At this stage, our results demonstrate that mutant CREBBP and EP300 proteins are deficient in acetylating BCL6 and p53, leading to constitutive activation of the oncoprotein and to decreased p53 tumour suppressor activity. The balance between the activities of these

two genes is critical for the regulation of DNA damage responses in mature germinal-centre cells during immunoglobulin gene remodeling^{36,44}. Thus, the consequences of BCL6 activity overriding p53 would be an increased tolerance for DNA damage in the context of diminished apoptotic and cell cycle arrest responses.

These results have important therapeutic implications in view of current attempts to use a variety of histone deacetylase inhibitors as anti-cancer drugs. Although the benefits of these compounds have been proved in certain malignancies, such as mature T-cell lymphoproliferative disorders, their efficacy in other common cancers, including B-NHL, is uncertain at this stage⁴⁵. The findings of this study suggest that the use of histone deacetylase inhibitors has a rational basis in B-NHL, because it may contribute to re-establishing physiologic acetylation levels. On the other hand, their efficacy should be re-evaluated by stratifying patients based on the presence of HAT defects as well as by testing the numerous histone deacetylase and sirtuin inhibitors with target specificity.

METHODS SUMMARY

Mutation analysis. The complete coding sequences and exon/intron junctions of *CREBBP* and *EP300* were analysed by PCR amplification and direct sequencing of whole-genome amplified DNA using the oligonucleotides reported in Supplementary Tables 5 and 6. Mutations were confirmed from both strands on independent PCR products amplified from genomic DNA, and their somatic origin was documented by analysis of matched normal DNA in available cases.

SNP array analysis. This was performed using Affymetrix Genome-Wide 6.0 Arrays and a computational workflow, which is described in detail in the Supplementary Information.

***In vivo* and *in vitro* characterization of HAT activity.** The ability of CREBBP mutants to acetylate BCL6 and p53 was assessed in human embryonic kidney (HEK293T) cells after co-transfection of pCMV-Flag-BCL6 (or pCIN4-Flag-p53) with plasmids encoding wild-type versus mutant HA-tagged mouse *Crebbp*. BCL6 acetylation was evaluated on Flag/M2 immunoprecipitates using antibodies directed against acetyl lysines; for p53, western blot analysis was performed on whole-cell extracts using a specific anti-acetylated p53 antibody. The amounts of exogenous CREBBP were monitored using anti-HA and anti-CREBBP (A22, Santa Cruz Biotechnology) antibodies. *In vitro* acetylation assays were performed using recombinant GST-p53 and purified CREBBP-HA proteins, in the presence of the indicated amounts of acetyl-CoA.

Transient transfection/reporter gene assays. The effect of CREBBP on BCL6-dependent transcription was assessed in HEK293T cells co-transfected with a luciferase reporter vector containing five BCL6 consensus binding sites upstream of the SV40 promoter (5XBCL6) and the pCMV-Flag-BCL6 construct, in the absence or presence of wild-type versus mutant CREBBP-HA expression vectors (see the online-only Methods).

Reconstitution of *Crebbp/Ep300* null MEFs. *Crebbp^{flox/flox};Ep300^{flox/flox};YFP⁺* conditional (double knockout) MEFs have been described³⁹. CREBBP expression was restored by retroviral infection with constructs encoding for HA-tagged wild-type or mutant *Crebbp*, and cells were analysed for H3K18 acetylation, cAMP-dependent transcriptional responses and cell proliferation (see the online-only Methods).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 June; accepted 2 December 2010.

1. Swerdlow, S. H. *et al.* WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (International Agency for Research on Cancer (IARC), Lyon, 2008).
2. Compagno, M. *et al.* Mutations of multiple genes cause deregulation of NF-κB in diffuse large B-cell lymphoma. *Nature* **459**, 717–721 (2009).
3. Davis, R. E. *et al.* Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* **463**, 88–92 (2010).
4. Klein, U. & Dalla-Favera, R. Germinal centres: role in B-cell physiology and malignancy. *Nature Rev. Immunol.* **8**, 22–33 (2008).
5. Lenz, G. *et al.* Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science* **319**, 1676–1679 (2008).
6. Lenz, G. & Staudt, L. M. Aggressive lymphomas. *N. Engl. J. Med.* **362**, 1417–1429 (2010).
7. Mandelbaum, J. BLIMP1 is a tumor suppressor gene frequently disrupted in activated B-cell like diffuse large B-cell lymphoma. *Cancer Cell* **18**, 568–579 (2010).

8. Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nature Genet.* **42**, 181–185 (2010).
9. Downing, J. R. Cancer genomes—continuing progress. *N. Engl. J. Med.* **361**, 1111–1112 (2009).
10. Goodman, R. H. & Smolik, S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **14**, 1553–1577 (2000).
11. Kalkhoven, E. CBP and p300: HATs for different occasions. *Biochem. Pharmacol.* **68**, 1145–1155 (2004).
12. Bannister, A. J. & Kouzarides, T. The CBP co-activator is a histone acetyltransferase. *Nature* **384**, 641–643 (1996).
13. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953–959 (1996).
14. Gu, W., Shi, X. L. & Roeder, R. G. Synergistic activation of transcription by CBP and p53. *Nature* **387**, 819–823 (1997).
15. Lill, N. L., Grossman, S. R., Ginsberg, D., DeCaprio, J. & Livingston, D. M. Binding and modulation of p53 by p300/CBP coactivators. *Nature* **387**, 823–827 (1997).
16. Avantaggiati, M. L. *et al.* Recruitment of p300/CBP in p53-dependent signal pathways. *Cell* **89**, 1175–1184 (1997).
17. Blobel, G. A., Nakajima, T., Eckner, R., Montminy, M. & Orkin, S. H. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proc. Natl Acad. Sci. USA* **95**, 2061–2066 (1998).
18. Bereshchenko, O. R., Gu, W. & Dalla-Favera, R. Acetylation inactivates the transcriptional repressor BCL6. *Nature Genet.* **32**, 606–613 (2002).
19. Grossman, S. R. *et al.* Polyubiquitination of p53 by a ubiquitin ligase activity of p300. *Science* **300**, 342–344 (2003).
20. Shi, D. *et al.* CBP and p300 are cytoplasmic E4 polyubiquitin ligases for p53. *Proc. Natl Acad. Sci. USA* **106**, 16275–16280 (2009).
21. Oike, Y. *et al.* Mice homozygous for a truncated form of CREB-binding protein exhibit defects in hematopoiesis and vasculo-angiogenesis. *Blood* **93**, 2771–2779 (1999).
22. Yao, T. P. *et al.* Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93**, 361–372 (1998).
23. Roelfsema, J. H. & Peters, D. J. Rubinstein-Taybi syndrome: clinical and molecular overview. *Expert Rev. Mol. Med.* **9**, 1–16 (2007).
24. Petrif, F. *et al.* Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature* **376**, 348–351 (1995).
25. Miller, R. W. & Rubinstein, J. H. Tumors in Rubinstein-Taybi syndrome. *Am. J. Med. Genet.* **56**, 112–115 (1995).
26. Iyer, N. G., Ozdag, H. & Caldas, C. p300/CBP and cancer. *Oncogene* **23**, 4225–4231 (2004).
27. Gayther, S. A. *et al.* Mutations truncating the EP300 acetylase in human cancers. *Nature Genet.* **24**, 300–303 (2000).
28. Ward, R., Johnson, M., Shridhar, V., van Deursen, J. & Couch, F. J. CBP truncating mutations in ovarian cancer. *J. Med. Genet.* **42**, 514–518 (2005).
29. Garbati, M. R., Alco, G. & Gilmore, T. D. Histone acetyltransferase p300 is a coactivator for transcription factor REL and is C-terminally truncated in the human diffuse large B-cell lymphoma cell line RC-K8. *Cancer Lett.* **291**, 237–245 (2010).
30. Shigeno, K. *et al.* Disease-related potential of mutations in transcriptional cofactors CREB-binding protein and p300 in leukemias. *Cancer Lett.* **213**, 11–20 (2004).
31. Borrow, J. *et al.* The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nature Genet.* **14**, 33–41 (1996).
32. Rowley, J. D. *et al.* All patients with the T(11;16)(q23;p13.3) that involves MLL and CBP have treatment-related hematologic disorders. *Blood* **90**, 535–541 (1997).
33. Sobulo, O. M. *et al.* MLL is fused to CBP, a histone acetyltransferase, in therapy-related acute myeloid leukemia with a t(11;16)(q23;p13.3). *Proc. Natl Acad. Sci. USA* **94**, 8732–8737 (1997).
34. Mullighan, C. G. *et al.* CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* doi: 10.1038/nature09727 (this issue).
35. Liu, X. *et al.* The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature* **451**, 846–850 (2008).
36. Phan, R. T. & Dalla-Favera, R. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. *Nature* **432**, 635–639 (2004).
37. Tang, Y., Zhao, W., Chen, Y., Zhao, Y. & Gu, W. Acetylation is indispensable for p53 activation. *Cell* **133**, 612–626 (2008).
38. Kwok, R. P. *et al.* Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature* **370**, 223–226 (1994).
39. Kasper, L. H. *et al.* CBP/p300 double null cells reveal effect of coactivator level and diversity on CREB transactivation. *EMBO J.* **29**, 3660–3672 (2010).
40. Bordoli, L. *et al.* Functional analysis of the p300 acetyltransferase domain: the PHD finger of p300 but not of CBP is dispensable for enzymatic activity. *Nucleic Acids Res.* **29**, 4462–4471 (2001).
41. Xu, W. *et al.* Global transcriptional coactivators CREB-binding protein and p300 are highly essential collectively but not individually in peripheral B cells. *Blood* **107**, 4407–4416 (2006).
42. Kung, A. L. *et al.* Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes Dev.* **14**, 272–277 (2000).
43. Legube, G. & Trouche, D. Regulating histone acetyltransferases and deacetylases. *EMBO Rep.* **4**, 944–947 (2003).
44. Phan, R. T., Saito, M., Basso, K., Niu, H. & Dalla-Favera, R. BCL6 interacts with the transcription factor Miz-1 to suppress the cyclin-dependent kinase inhibitor p21 and cell cycle arrest in germinal center B cells. *Nature Immunol.* **6**, 1054–1060 (2005).
45. Stimson, L., Wood, V., Khan, O., Fotheringham, S. & La Thangue, N. B. HDAC inhibitor-based therapies and haematological malignancy. *Ann. Oncol.* **20**, 1293–1302 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank Q. Shen and the Molecular Pathology Shared Resource of the Herbert Irving Comprehensive Cancer Center at Columbia University for histology service; W. Gu for discussions and reagents; Y. Tang, M. Li and D. Chao for suggestions; V. Bardwell for the Bcl6 reporter construct; and R. S. K. Chaganti for sharing unpublished information. Whole-exome capture and sequencing were conducted at Roche NimbleGen and 454 Life Sciences. Automated DNA sequencing was performed at Genewiz Inc. This work was supported by NIH grants P01-CA092625 and R01-CA37295 (to R.D.-F.), a Specialized Center of Research grant from the Leukemia and Lymphoma Society (to R.D.-F.), NIH grant DE018183, a Cancer Center (CORE) support grant P30 CA021765, and the American Lebanese Syrian Associated Charities of St Jude Children's Research Hospital (to P.K.B.), the Northeast Biodefence Center (U54-AI057158) and the National Library of Medicine (1R01LM010140-01) (to R.R.), and the AIRC Special Program Molecular Clinical Oncology—5 per mille (contract number 10007, Milan) (to G.G.). A. Chiarenza is on leave from the Division of Hematology, Ospedale Ferrarotto, University of Catania, Catania, Italy. L.P. is on leave from the University of Perugia Medical School, Perugia, Italy.

Author Contributions L.P. and R.D.-F. designed the study and wrote the manuscript, with contributions from all authors. L.P. designed and conducted experiments, analysed data and coordinated the study. D.D.-S. designed and conducted experiments, and analysed immunohistochemistry data. A. Chiarenza, G.F. and A.G. conducted CREBBP/EP300 amplification and sequencing analysis. L.H.K., S.L. and P.K.B. were responsible for the experiments in MEF cells. H.T. performed immunohistochemistry and immunofluorescence staining of human tissue biopsies. V.V.M. developed FISH assays and analysed cytogenetic data. C.G.M. and J.M. analysed microarray data. A. Chadburn, D.R. and G.G. provided well-characterized patient samples. V.T. and R.R. developed algorithms and analysed high-throughput sequencing data.

Author Information The Affymetrix expression data reported in this paper have been deposited in the NCBI Gene Expression Omnibus (GEO) database (Series Accession Number GSE12195). The SNP Array 6.0 data and the whole exome sequencing data from the seven DLBCL cases have been deposited in dbGaP under accession number phs000328.v1.p1. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to L.P. (lp171@columbia.edu) or R.D.-F. (rd10@columbia.edu).

METHODS

DNA extraction, amplification and sequencing. Genomic DNA was extracted by standard methods, and whole-genome amplification was performed using the RepliG kit (Qiagen) according to the manufacturer's instructions. Sequences for all annotated exons and flanking introns of *CREBBP* and *EP300* were obtained from the UCSC Human Genome database, using the corresponding mRNA accession number as a reference (NM_004380.2 and NM_001429.3, respectively). PCR primers, located ≥ 50 bp upstream or downstream to target exon boundaries, were designed in the Primer 3 program (<http://frodo.wi.mit.edu/primer3/>) and filtered using UCSC In Silico PCR (<http://genome.ucsc.edu>) to exclude pairs yielding more than a single product (Supplementary Tables 5 and 6). Purified amplicons were sequenced directly from both strands as described², and compared to the corresponding germline sequences, using the Mutation Surveyor Version 2.41 software package (Softgenetics; <http://www.softgenetics.com>).

Somatic mutations were confirmed on independent PCR products obtained from high-molecular-weight genomic DNA. Synonymous mutations, previously reported polymorphisms (Human dbSNP Database at NCBI, Build 130, and Ensembl Database) and changes present in the matched normal DNA, when available, were excluded. In cases carrying multiple events within a single gene, the allelic distribution of the mutations was determined by cloning and sequencing PCR products obtained from cDNA and spanning both events ($n = 10$ clones each).

Northern blot analysis of CREBBP and EP300 expression. Total RNA (12 μ g) was extracted from exponentially growing cell lines by TRIzol (Invitrogen) and Northern blot analysis was performed according to standard procedures, with radiolabelled probes corresponding to a 0.9-kb fragment of the human *CREBBP* cDNA (region 204–1143, GenBank accession number NM_004380.2) or a 1.2-kb fragment of the human *EP300* cDNA (region 6277–7524, GenBank accession number NM_001429.3), and GAPDH as a control for loading. *CREBBP* and *EP300* expression levels were then quantitated by Phosphorimager analysis and normalized with the GAPDH levels.

Tissue microarrays, immunohistochemistry and immunofluorescence analysis. DLBCL tissue microarrays were constructed according to standard procedures and analysed by immunohistochemistry, using rabbit polyclonal antibodies directed against the N-terminus of CREBBP (A22) or EP300 (N15) (Santa Cruz Biotechnology). Cases were scored as positive if $\geq 20\%$ tumour cells were stained by the antibody. Immunofluorescence analysis of CREBBP expression in conditional *Crebbp/Ep300* double-knockout MEFs was performed three days after deletion of the endogenous loci and five days after infection with HA-tagged CREBBP retrovirus, using the CREBBP antibody (A22). The specificity of both the A22 and the N15 antibodies had been previously validated on paraffin-embedded cell pellets from HEK293T cells transfected with control and CREBBP-specific short hairpin RNAs as well as by immunofluorescence staining of *Crebbp/Ep300* double-knockout MEFs (data not shown).

Protein extraction and western blot analysis. Whole-cell extracts were prepared in RIPA buffer containing protease inhibitors as described¹⁸ and were analysed by western blotting according to standard methods, using the following primary antibodies: anti-FLAG/M2 (Sigma), anti-HA (3F10) (Roche), anti-BCL6 (GI191E/A8, Cell Marque), rabbit anti-acetyl lysines (Cell Signaling), anti-acetylated p53 (kind gift of W. Gu), anti-p53 (DO1), anti-CREBBP (A22), anti-EP300 (N15) (all from Santa Cruz Biotechnology), anti-GFP (JL-8) (Clontech), anti- β -actin (clone AC-15) and anti- β -tubulin (clone B-5-1-2) (both from Sigma). Proteins were resolved by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) in 3–8% NuPAGE Tris-Acetate gels (Invitrogen), and visualized using a chemiluminescence detection kit (Pierce) as recommended by the manufacturer.

Transient transfections/reporter gene assays. Transient transfections were performed in HEK293T cells using polyethylenimine, as described⁴⁶. For reporter assays, cells were seeded on a 24-well plate and transfected using 100 ng of a luciferase reporter construct driven by a PGL3-SV40-based backbone with five BCL6 binding sites⁴⁷, 2.5 ng of pRL-SV40 (Promega), 1 ng of BCL6-encoding plasmid and the indicated doses of wild-type and mutant CREBBP expression vectors. The total amount of transfected DNA was kept constant in each experiment by adding pCMV–HA or pcDNA3 vector sequences to a final amount of 600 ng per well, and 3.5 μ g of polyethylenimine. All experiments were performed in duplicate and luciferase activities were measured 36 h after transfection using the Dual-Luciferase Reporter Assay System (Promega), according to the manufacturer's instructions. Given the known effects of CREBBP on the SV40 promoter and on other regulatory sequences, the response of the BCL6 reporter was normalized first to that of a SV40–Renilla reporter construct, and then to the basal activity of CREBBP on the 5X-BCL6 reporter, in the absence of BCL6. Data are

expressed as differences relative to the basal activity of the reporter construct (set to 1) after the above-mentioned normalization.

Co-immunoprecipitation assays. To assess the interaction between CREBBP and BCL6 or p53, HEK293T cells were co-transfected with plasmids encoding for the two proteins, together with wild-type or mutant CREBBP–HA. Thirty-six hours after transfection (or twenty-four hours for p53), cells were lysed in immunoprecipitation buffer (50 mM Tris, pH 7.0, 250 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.05% NP40, 10 mM sodium fluoride, 0.1 mM sodium orthovanadate and protease inhibitor cocktail) (from Sigma), and the cleared lysates were incubated overnight at 4 °C with anti-HA or anti-Flag/M2 beads (Sigma). Immunocomplexes were eluted by incubating the beads in immunoprecipitation buffer containing HA or 3X-Flag peptide, respectively (Sigma). A fraction of the final eluates was resolved by SDS–PAGE and analysed by western blot.

In vitro acetyl-transferase assays. Recombinant GST–p53 and CREBBP–HA proteins were obtained as described in detail in the Supplementary Information and used for *in vitro* acetylation assays according to published protocols, with minor modifications^{14,48,49}. Briefly, 10–24 ng of recombinant CREBBP–HA and 100 ng of GST–p53 were combined in 40 μ l reactions containing 50 mM Tris–Cl (pH 8.0), 10% glycerol, 1 mM DTT, 1 mM PMSF, 0.1 mM EDTA, 10 mM butyric acid (Sigma), and variable amounts of acetyl–CoA (Sigma) (2 mM to 25 nM). As shown in Supplementary Fig. 6, acetyl–CoA concentrations as low as 25 nM were sufficient to obtain efficient acetylation of the substrate by CREBBP. Reactions were performed for 60 min at 30 °C, and stopped by addition of an equal volume of 2 \times Laemmli buffer followed by heating for 10 min at 70 °C. A fraction of the final product was resolved by SDS–PAGE in 3–8% Tris–Acetate gradient gels (Invitrogen) and analysed by western blotting.

Reconstitution and analysis of conditional *Crebbp/Ep300* double-knockout MEFs. The protocols for MEF isolation, cell culture and retroviral transduction, as well as the mouse *Crebbp* (CBP)–HA retroviral construct used as the template for the generation of various CREBBP mutants have been described previously⁵⁰. In this system, Cre-mediated recombination induces expression of YFP, allowing for specific identification of the deleted cells. *Crebbp*^{flax/flax}; *Ep300*^{flax/flax}; YFP⁺ MEFs were first infected with retroviruses encoding either wild-type or selected CREBBP point mutants, and the endogenous *Crebbp*^{flax} and *Ep300*^{flax} loci were deleted after 48 h by infection with Cre-expressing adenovirus (Ad-Cre). The W1502A/Y1503S HAT-dead mutation was used as negative control⁴⁰. In all experiments, transduction efficiencies were 70% or higher, as assessed by immunofluorescence analysis of HA-positive cells three days after deletion of endogenous *Crebbp/Ep300* and five days after retroviral infection. To examine cAMP-dependent gene expression, MEFs were cultured for 16 h in DMEM containing 0.1% FBS and treated for 90 min with 10 μ M forskolin + 100 μ M IBMX (or ethanol vehicle) before harvesting in TRIzol (Invitrogen); quantitative RT–PCR assays were performed as reported⁵⁰. Expression of *Crebbp* and *Ep300* was verified by immunofluorescence analysis using rabbit polyclonal anti-CREBBP (A22) and anti-EP300 (N20) antibodies (Santa Cruz Biotechnology). Analysis of H3K18 acetylation was performed as described, using the anti-H3K18Ac antibody (ab1191) (Abcam) and the HA-11 monoclonal antibody against the HA epitope (Boehringer Mannheim). Confocal images were taken using the same settings for all mutants and mean intensity ratios for individual nuclei were collected using SlideBook 5 software. Nuclei in which the *Crebbp*–HA signal was at least 2.5-fold above background were used to calculate the ratio of the H3K18Ac mean signal intensity to the *Crebbp*–HA mean signal intensity. For the growth assays shown in Supplementary Fig. 11, *Crebbp*^{flax/flax}; *Ep300*^{flax/flax}; YFP⁺ MEFs were infected with CREBBP retroviruses 48 h before Ad-Cre-mediated deletion of endogenous *Crebbp* and *Ep300*. From each reconstituted population, equivalent numbers of YFP⁺ (double-knockout) MEFs were seeded at day 1 (that is, the day following overnight Ad-Cre treatment), and the total number of YFP⁺ cells was calculated on day 11 from the total cell number, based on the percentage of YFP⁺ cells as assessed by flow cytometric analysis.

46. Bieber, T. & Elsas, H. P. Preparation of a low molecular weight polyethylenimine for efficient cell transfection. *Biotechniques* **30**, 74–77, 80–81 (2001).
47. Huynh, K. D., Fischle, W., Verdine, E. & Bardwell, V. J. BCoR, a novel corepressor involved in BCL-6 repression. *Genes Dev.* **14**, 1810–1823 (2000).
48. Kuninger, D., Lundblad, J., Semirale, A. & Rotwein, P. A non-isotopic *in vitro* assay for histone acetylation. *J. Biotechnol.* **131**, 253–260 (2007).
49. Tang, Y., Luo, J., Zhang, W. & Gu, W. Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. *Mol. Cell* **24**, 827–839 (2006).
50. Bedford, D. C., Kasper, L. H., Fukuyama, T. & Brindle, P. K. Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics* **5**, 9–15 (2010).

Coupled quantized mechanical oscillators

K. R. Brown¹, C. Ospelkaus¹, Y. Colombe¹, A. C. Wilson¹, D. Leibfried¹ & D. J. Wineland¹

The harmonic oscillator is one of the simplest physical systems but also one of the most fundamental. It is ubiquitous in nature, often serving as an approximation for a more complicated system or as a building block in larger models. Realizations of harmonic oscillators in the quantum regime include electromagnetic fields in a cavity^{1–3} and the mechanical modes of a trapped atom⁴ or macroscopic solid⁵. Quantized interaction between two motional modes of an individual trapped ion has been achieved by coupling through optical fields⁶, and entangled motion of two ions in separate locations has been accomplished indirectly through their internal states⁷. However, direct controllable coupling between quantized mechanical oscillators held in separate locations has not been realized previously. Here we implement such coupling through the mutual Coulomb interaction of two ions held in trapping potentials separated by 40 μm (similar work is reported in a related paper⁸). By tuning the confining wells into resonance, energy is exchanged between the ions at the quantum level, establishing that direct coherent motional coupling is possible for separately trapped ions. The system demonstrates a building block for quantum information processing and quantum simulation. More broadly, this work is a natural precursor to experiments in hybrid quantum systems, such as coupling a trapped ion to a quantized macroscopic mechanical or electrical oscillator^{9–13}.

The direct coupling of atomic ions in separate potential wells is a key feature of proposals to implement quantum simulation^{14–16}, and it could allow logic operations to be performed in a multi-zone quantum information processor^{10,17,18} without the requirement of bringing the ion qubits into the same trapping potential. Moreover, the coupling could prove useful for metrology and sensing. For example, it could extend the capabilities of quantum logic spectroscopy^{9,19,20} to ions that cannot be trapped within the same potential well as the measurement ion, such as oppositely charged ions or even antimatter particles^{9,10}. Coupling could be obtained either through mutually shared electrodes^{9,21} or directly through the Coulomb interaction^{10,17,22,23}.

The Coulomb interaction potential for two trapped ions, a and b , with charges q_a and q_b in potential wells separated by a distance s_0 is given by

$$U(x_a, x_b) = \frac{1}{4\pi\epsilon_0 s_0} \frac{q_a q_b}{-x_a + x_b} \approx \frac{1}{4\pi\epsilon_0 s_0} \left(1 + \frac{x_a - x_b}{s_0} + \frac{x_a^2}{s_0^2} + \frac{x_b^2}{s_0^2} - \frac{2x_a x_b}{s_0^2} \right)$$

Here x_a and x_b are the displacements of the ions from the external potential minima and ϵ_0 is the permittivity of free space. The first term is constant and does not affect the dynamics. The second term represents a steady force between the ions that displaces them slightly; if necessary, it can be counteracted with additional potentials applied to nearby electrodes. The terms proportional to x_a^2 and x_b^2 represent static changes in the trap frequencies that could also be compensated with potentials applied to nearby electrodes. The term proportional to $x_a x_b$ represents the lowest-order coupling between the ions' motions. For small deviations, x'_a and x'_b , from equilibrium, the coupling is

$$\frac{-q_a q_b}{2\pi\epsilon_0 s_0^3} (x'_a x'_b) = -\hbar\Omega_{\text{ex}}(a + a^\dagger)(b + b^\dagger) \approx -\hbar\Omega_{\text{ex}}(ab^\dagger + a^\dagger b) \quad (1)$$

where

$$\Omega_{\text{ex}} \equiv \frac{q_a q_b}{4\pi\epsilon_0 s_0^3 \sqrt{m_a m_b} \sqrt{\omega_{0a} \omega_{0b}}} \quad (2)$$

and a , a^\dagger , b and b^\dagger represent the harmonic oscillator lowering (a , b) and raising (a^\dagger , b^\dagger) operators, m_i and ω_{0i} are respectively the ion masses and motional frequencies, \hbar is Planck's constant divided by 2π , and we have neglected fast-rotating terms. Minimizing the distance, s_0 , between ions is crucial, because for fixed values of ω_{0i} , the coupling rate scales as $\Omega_{\text{ex}} \propto 1/s_0^3$.

When $\omega_{0a} = \omega_{0b} = \omega_0$ (the resonance condition), we find that

$$\begin{aligned} a^\dagger(t) &= \exp(i\omega_0 t) (a^\dagger(0) \cos(\Omega_{\text{ex}} t) - ib^\dagger(0) \sin(\Omega_{\text{ex}} t)) \\ b^\dagger(t) &= \exp(i\omega_0 t) (b^\dagger(0) \cos(\Omega_{\text{ex}} t) - ia^\dagger(0) \sin(\Omega_{\text{ex}} t)) \end{aligned} \quad (3)$$

At time $t = \tau_{\text{ex}} \equiv \pi/2\Omega_{\text{ex}}$, the operators have changed roles up to a phase factor, and the oscillators have completely swapped their energies, regardless of their initial states. At $t = 2\tau_{\text{ex}}$, the energies have returned to their initial values in each ion. The mean occupation, $\langle a^\dagger a \rangle$, of ion a as a function of time exhibits oscillations with period $2\tau_{\text{ex}}$.

Figure 1 shows a micrograph of our surface-electrode trap²⁴, constructed of gold electrodes, 8 μm thick with 5- μm gaps, electroplated

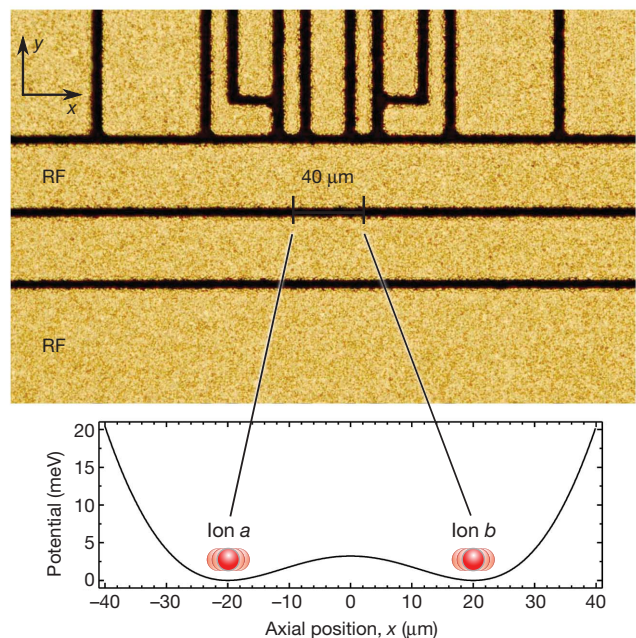


Figure 1 | Micrograph of the ion trap, showing radio-frequency (RF) and d.c. electrodes, and gaps between electrodes (darker areas). The lower part of the figure indicates the simulated potential along the trap x axis. Two trapping wells are separated by 40 μm , with ion positions marked by red spheres. The d.c. electrodes are sufficient to control the axial frequency and the position of each ion independently. Here both frequencies are ~ 4 MHz and the potential barrier between the two ions is ~ 3 meV.

¹Time and Frequency Division, National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA.

onto a crystalline quartz substrate. The trap can produce two potential minima at a height $d_0 = 40 \mu\text{m}$ above the surface and separated by $s_0 = 40 \mu\text{m}$ along the x axis. Each potential well confines a single $^9\text{Be}^+$ ion with an axial (parallel to x) motional frequency of $\omega_0/2\pi \approx 4 \text{ MHz}$ and a barrier between wells of $\sim 3 \text{ meV}$. Pseudopotential confinement in radial directions (normal to x) is accomplished with a peak potential of $\sim 100 \text{ V}$ at 170 MHz applied to the radio-frequency electrodes, yielding radial frequencies of $\sim 22 \text{ MHz}$. By applying static potentials to the d.c. electrodes, we can independently vary the separation between the ions and the curvatures of the two trapping wells. In this way, the ion axial motional frequencies can be brought into or out of resonance, allowing a tunable interaction. For $\omega_0/2\pi = 4.04 \text{ MHz}$, we predict that $\tau_{\text{ex}} = 162 \mu\text{s}$, where we have included a 2% correction in Ω_{ex} owing to the metallic electrodes beneath the ions (Methods Summary).

With currently achieved size scales in ion traps, the Coulomb interaction is relatively weak, so low ion heating rates and stable trapping potentials are essential. Heating rates can be suppressed by operating at cryogenic temperatures^{25,26}, such that the direct Coulomb coupling rate can exceed the heating rate. In this work, the trap electrodes and surrounding vacuum enclosure are cooled to 4.2 K with a liquid-helium bath cryostat. With similar versions of this trap at $\omega_0/2\pi = 2.3 \text{ MHz}$, heating rates expressed as $d\langle n \rangle/dt$ (where n denotes the quantum number of motional Fock state $|n\rangle$ and $\langle n \rangle$ is its expectation value) were observed to be as low as 70 quanta per second, consistent with the results of ref. 26. However, for the experiments described here ($\omega_0/2\pi \approx 4.0\text{--}5.6 \text{ MHz}$) the heating rate was $\sim 500\text{--}2,000$ quanta per second and varied between the two wells. We observed $d\langle n \rangle/dt \propto 1/\omega_0^2$ in this trap, in agreement with previous reports^{25–27}, so large values of ω_0 are beneficial. The use of $^9\text{Be}^+$, the lightest of the commonly trapped atomic ions, is an advantage here, because for given d.c. trapping potentials the heating rate should remain unchanged while $\Omega_{\text{ex}} \propto m^{-1/2}$. Cryogenic operation decreases the background gas pressure to negligible levels, such that ion loss rates due to collisions with background gas are smaller than one per day.

A signature of coupling between the ions is the splitting between the two axial normal mode frequencies. As the trap potential is tuned into the resonance condition, this splitting, δf , reaches a theoretical minimum of $\delta f = \Omega_{\text{ex}}/\pi = 3.1 \text{ kHz}$. A plot of the mode frequencies will therefore show an avoided crossing. We measure the mode frequencies by applying a nearly resonant oscillating potential pulse to one of the trap electrodes. We then illuminate both ions with laser radiation resonant with the $^2S_{1/2} \rightarrow ^2P_{3/2}$ cycling transition at 313 nm . A decrease in the resulting fluorescence indicates that a mode of the ions' motion has been resonantly excited. For pulse lengths $\tau_p \gg 1/\delta f$, we resolve the two modes (Fig. 2a). We sweep the trapping wells through resonance by varying the static potentials that are applied to the trap electrodes. A plot of the resulting mode frequencies, determined as above, is given in Fig. 2b, c, showing a minimum of $\delta f = 3.0(5) \text{ kHz}$, in agreement with theory.

To demonstrate coupling at the level of a few motional quanta, we observe the exchange of energy between the two ions as follows. The ion motional frequencies are initially detuned by 100 kHz , which is much greater than $\Omega_{\text{ex}}/2\pi$, effectively decoupling the ions' motions. The ions are then simultaneously illuminated with a laser detuned by -10 MHz from the $^2S_{1/2} \rightarrow ^2P_{3/2}$ cycling transition, cooling them into a thermal state at the Doppler limit with mean occupation $\langle n \rangle = 2.3(1)$. Subsequently, ion a is cooled to $\langle n_a \rangle = 0.35(2)$ by several cycles of stimulated Raman cooling with the $|\downarrow\rangle \equiv |F=2, m_F=-2\rangle$ and $|\uparrow\rangle \equiv |1, -1\rangle$ hyperfine states²⁸. The Raman beams are counter-propagating and oriented at 45° relative to the x axis. At this point, the potentials are brought into resonance ($\omega_0/2\pi = 4.04 \text{ MHz}$) within an interval ($9 \mu\text{s}$) short in comparison with τ_{ex} but long in comparison with the axial oscillation period. They remain on resonance for a time τ , allowing energy to transfer between the ions. After a time τ , the potentials are adiabatically returned to their off-resonance values and we determine the mean quantum number, $\langle n_a \rangle$, in ion a by observing asymmetry between the red and blue motional sidebands of the $|\downarrow\rangle$ -to- $|\uparrow\rangle$ hyperfine Raman transition²⁸.

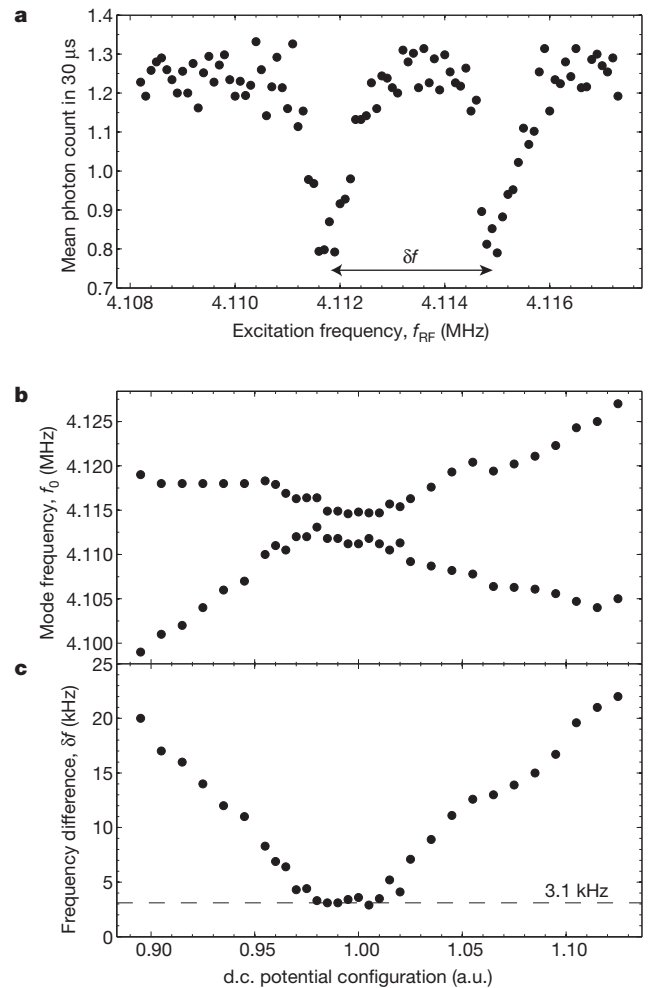


Figure 2 | Motional spectroscopy of two coupled ions near the avoided crossing. **a**, Decreases in collected fluorescence occur at values of excitation frequency, f_{RF} , corresponding to the ion mode frequencies. With $\tau_p = 960 \mu\text{s}$, the splitting on resonance is resolved. **b**, **c**, Mode frequencies (**b**) and mode frequency splitting, δf (**c**), for the axial normal modes of two ions separated by $40 \mu\text{m}$. Error bars are smaller than the size of the points. The data were acquired over a 1-h period, and slow variations in ambient potentials gave rise to the fluctuations.

As seen in Fig. 3, energy exchanges between the ions during an interval $\tau_{\text{ex}} = 155(1) \mu\text{s}$. The 5% disagreement between the measured and predicted ($162 \mu\text{s}$) values for τ_{ex} is probably due to uncertainty in the ion separation, s_0 (even a $1\text{-}\mu\text{m}$ uncertainty would account for the disagreement). The first maximum of $\langle n_a \rangle$ corresponds to the cooling limit of ion b ($\langle n_b(\tau=0) \rangle = 2.3(1)$ quanta). The underlying linear growth in $\langle n_a \rangle$ corresponds to heating of the ions at a rate of $\dot{n} = 1,885(10)$ quanta per second (Methods Summary).

As a final experiment, we demonstrate energy exchange at approximately the single-quantum level. Ideally, the experiment takes the following form. The ions are tuned to the resonance condition throughout and are initially Doppler-cooled. Ion a is Raman-cooled, sympathetically cooling ion b and thereby preparing the state $|0\rangle_a |\downarrow\rangle_a |0\rangle_b$. To create a single motional quantum, we drive ion a with a blue-sideband Raman π pulse (of duration $10 \mu\text{s}$, which is much less than τ_{ex}), creating the state $|1\rangle_a |\uparrow\rangle_a |0\rangle_b$. The system oscillates between $|1\rangle_a |\uparrow\rangle_a |0\rangle_b$ and $|0\rangle_a |\uparrow\rangle_a |1\rangle_b$ with period $2\tau_{\text{ex}}$. After a time τ , we drive ion a with another blue-sideband π pulse, conditionally flipping the spin from $|1\rangle_a |\uparrow\rangle_a$ to $|0\rangle_a |\downarrow\rangle_a$, dependent on the presence of a motional quantum in ion a . The final internal state probability will be given by $P(|\uparrow\rangle_a)(\tau) = \sin^2(\Omega_{\text{ex}}\tau)$. In practice, contrast in the oscillations (Fig. 4) is significantly reduced by

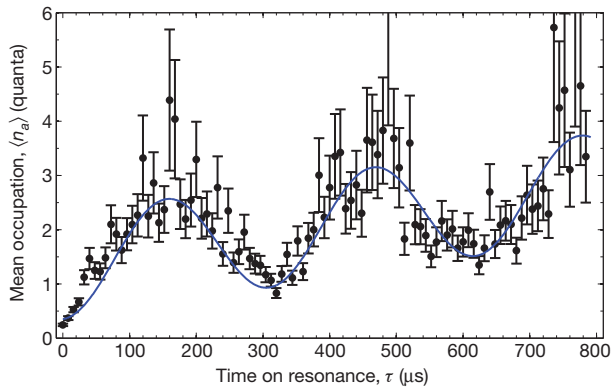


Figure 3 | Energy swapping between two ions in separate trapping potentials at the level of a few quanta. The mean occupation, $\langle n_a \rangle$, of ion a is plotted with error bars (s.e.m.) for various durations, τ , that the ion motional frequencies remain on resonance. The blue curve represents a fit to theory with four free parameters: the two initial mean quantum numbers, the exchange time and the heating rate. Energy exchanges between the ions at 155(1)- μ s intervals. The linearly increasing trend in $\langle n_a \rangle$ is due to ion heating at a rate of 1,885(10) quanta per second. Uncertainties represent standard errors of the fit parameters.

incomplete cooling, motional decoherence and decoherence due to imperfect Raman sideband pulses. We estimate that ion a is cooled initially to $\langle n_a \rangle = 0.3(1)$. Although we were unable to measure the initial temperature of ion b directly, comparison of the contrast and temporal behaviour of our exchange data (Fig. 4) with simulations indicates that $\langle n_b \rangle \lesssim 0.6$. Motional decoherence results from heating and from trap frequency instability over the time required to acquire the data. Raman sideband pulses suffer from variations in laser intensity and fluctuations in the sideband coupling caused by thermal spread in the y and z motional states (Debye–Waller factors¹⁰). For $\omega_0/2\pi = 5.56$ MHz, we observe oscillations with period $2\tau_{\text{ex}} = 437(4)$ μ s. The 2% disagreement with the prediction (447 μ s) of equations (2) and (3) is probably due to uncertainty in the ion separation and the difficulty of maintaining the exact resonance condition.

Significant improvements in coupling fidelity seem to be in reach with current technology. It should be possible to improve Raman laser intensity stability, and Debye–Waller factors from the y and z motion can be eliminated by proper choice of beam directions²⁸. We believe that our motional frequency instability can be alleviated by ensuring that the trap surface is free of charged contaminants. Faster exchange can be achieved by scaling down trap dimensions, but this puts a premium on reduced motional heating. Although the work presented

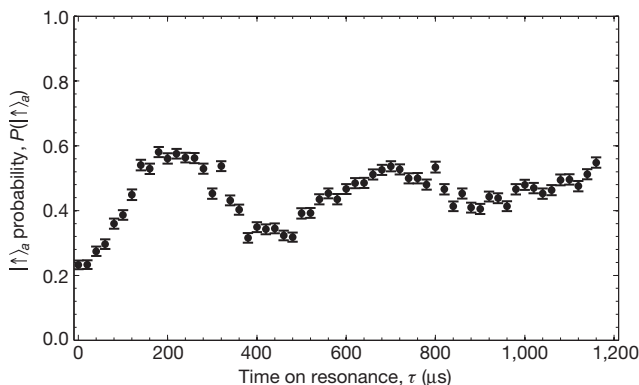


Figure 4 | Motional exchange between two ions in separate trapping potentials at approximately the single-quantum level. The probability, $P(|\uparrow\rangle_a)$, of measuring ion a in spin state $|\uparrow\rangle_a$ at the end of the experimental sequence is plotted with error bars (s.e.m.) against the time, τ , for which the ions interact. $P(|\uparrow\rangle_a)$ oscillates with period $2\tau_{\text{ex}} = 437(4)$ μ s as a quantum exchanges between the ions.

here uses the axial mode to couple the ions, it may prove advantageous to use the radial modes in future experiments because of the lower heating rates associated with their higher frequencies. As an example, radial mode frequencies of ~ 30 MHz are routinely achieved in the apparatus, compared with axial frequencies of $\lesssim 10$ MHz.

These results could lead to several possible applications in quantum state engineering and spectroscopy. For example, from the motional state $|1\rangle_a|0\rangle_b$ the state of the ions at time $t = \tau_{\text{ex}}/2$ is the Bell state $(|1\rangle_a|0\rangle_b + i|0\rangle_a|1\rangle_b)/\sqrt{2}$. Transferring this state onto the ions' internal states with sideband pulses would create an entangled spin state, even between ions of dissimilar species. This could be used as an entangled-pair factory in the scheme of refs 10,18, with the advantage over previous schemes⁷ that the ions are already in separate wells, ready for distribution to separate locations. The coupling could be used to read out the state of one ion species with another, an ability useful for error correction protocols and for quantum logic spectroscopy¹⁹. When the harmonic wells are not in resonance, the spin state of one ion can be read out without destroying the state of the other, so schemes for weak or quantum non-demolition measurements that use either Kerr-type nonlinearities or quantum logic may become feasible¹. Hybrid quantum systems, incorporating similar interactions to couple a trapped ion to other quantum devices, could serve as a means of transferring quantum information between different qubit implementations in a future quantum network. For example, a trapped ion could act as a quantum transformer between a superconducting qubit¹³ and a photonic qubit^{29,30}. The sympathetic cooling through exchange might also be used to cool neutral molecules³¹.

METHODS SUMMARY

The shielding factor, β , represents the ratio of the exchange rates Ω_{ex} with and without the presence of the trapping electrodes. To a good approximation, we ensure that at the motional frequencies all trap electrodes are held at ground. Therefore, assuming that gaps between the electrodes are negligible, the shielding factor can be calculated with the method of images. The result is

$$\beta = 1 - \frac{1}{2} \left(\frac{3(s_0/d_0)^5}{(4 + (s_0/d_0)^2)^{5/2}} - \frac{(s_0/d_0)^3}{(4 + (s_0/d_0)^2)^{3/2}} \right)$$

which reaches a maximum of $\beta = 1.018$ at $s_0 = d_0$.

The evolution of $\langle n_a \rangle$ under the influence of equation (1), including heating effects and assuming that both ions begin in a thermal state with mean quantum numbers n_{a0} and n_{b0} , can be predicted with a Langevin equation⁹. On resonance, the evolution is

$$\langle n_a \rangle = n_{a0} \cos^2(\Omega_{\text{ex}} t) + n_{b0} \sin^2(\Omega_{\text{ex}} t) + \dot{n}$$

where \dot{n} represents the mean of $d\langle n_a \rangle/dt$ and $d\langle n_b \rangle/dt$ for uncorrelated noise sources.

Received 9 November; accepted 1 December 2010.

Published online 23 February 2011.

1. Haroche, S. & Raimond, J.-M. *Exploring the Quantum: Atoms, Cavities, and Photons* (Oxford Univ. Press, 2006).
2. Miller, R. *et al.* Trapped atoms in cavity QED: coupling quantized light and matter. *J. Phys. B* **38**, S551 (2005).
3. Houck, A. A. *et al.* Generating single microwave photons in a circuit. *Nature* **449**, 328–331 (2007).
4. Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. J. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75**, 281–324 (2003).
5. O'Connell, A. D. *et al.* Quantum ground state and single-phonon control of a mechanical resonator. *Nature* **464**, 697–703 (2010).
6. Monroe, C. *et al.* in *Atomic Physics 17* (eds Arimondo, E., De Natale, P. & Inguscio, M.) 173–186 (Proc. 17th Int. Conf. Atomic Phys., Springer, 2001).
7. Jost, J. D. *et al.* Entangled mechanical oscillators. *Nature* **459**, 683–685 (2009).
8. Harlander, M., Lechner, R., Brownnutt, M., Blatt, R. & Hänsel, W. Trapped-ion antennae for the transmission of quantum information. *Nature* doi:10.1038/nature09800 (this issue).
9. Heinzen, D. J. & Wineland, D. J. Quantum-limited cooling and detection of radio-frequency oscillations by laser-cooled ions. *Phys. Rev. A* **42**, 2977–2994 (1990).
10. Wineland, D. J. *et al.* Experimental issues in coherent quantum-state manipulation of trapped atomic ions. *J. Res. Natl Inst. Stand. Technol.* **103**, 259–328 (1998).
11. Tian, L. & Zoller, P. Coupled ion-nanomechanical systems. *Phys. Rev. Lett.* **93**, 266403 (2004).

12. Hensinger, W. K. *et al.* Ion trap transducers for quantum electromechanical oscillators. *Phys. Rev. A* **72**, 041405(R) (2005).
13. Tian, L., Blatt, R. & Zoller, P. Scalable ion trap quantum computing without moving ions. *Eur. Phys. J. D* **32**, 201–208 (2005).
14. Schmied, R., Roscilde, T., Murg, V., Porras, D. & Cirac, J. I. Quantum phases of trapped ions in an optical lattice. *N. J. Phys.* **10**, 045017 (2008).
15. Chiaverini, J. & Lybarger, W. E. Jr. Laserless trapped-ion quantum simulations without spontaneous scattering using microtrap arrays. *Phys. Rev. A* **77**, 022324 (2008).
16. Schmied, R., Wesenberg, J. H. & Leibfried, D. Optimal surface-electrode trap lattices for quantum simulation with trapped ions. *Phys. Rev. Lett.* **102**, 233002 (2009).
17. Cirac, J. I. & Zoller, P. A scalable quantum computer with ions in an array of microtraps. *Nature* **404**, 579–581 (2000).
18. Kielpinski, D., Monroe, C. & Wineland, D. J. Architecture for a large-scale ion-trap quantum computer. *Nature* **417**, 709–711 (2002).
19. Schmidt, P. O. *et al.* Spectroscopy using quantum logic. *Science* **309**, 749–752 (2005).
20. Rosenband, T. *et al.* Frequency ratio of Al^+ and Hg^+ single-ion optical clocks; metrology at the 17th decimal place. *Science* **319**, 1808–1812 (2008).
21. Daniilidis, N., Lee, T., Clark, R., Narayanan, S. & Häffner, H. Wiring up trapped ions to study aspects of quantum information. *J. Phys. B* **42**, 154012 (2009).
22. Tan, J. N. Interacting ion oscillators in contiguous confinement wells. *Bull. Am. Phys. Soc.* **47**, 103 (2002).
23. Ciaramicoli, G., Marzoli, I. & Tombesi, P. Scalable quantum processor with trapped electrons. *Phys. Rev. Lett.* **91**, 017901 (2003).
24. Seidelin, S. *et al.* Microfabricated surface-electrode ion trap for scalable quantum information processing. *Phys. Rev. Lett.* **96**, 253003 (2006).
25. Deslauriers, L. *et al.* Scaling and suppression of anomalous heating in ion traps. *Phys. Rev. Lett.* **97**, 103007 (2006).
26. Labaziewicz, J. *et al.* Temperature dependence of electric field noise above gold surfaces. *Phys. Rev. Lett.* **101**, 180602 (2008).
27. Epstein, R. J. *et al.* Simplified motional heating rate measurements of trapped ions. *Phys. Rev. A* **76**, 033411 (2007).
28. Monroe, C. *et al.* Resolved-sideband Raman cooling of a bound atom to the 3D zero-point energy. *Phys. Rev. Lett.* **75**, 4011–4014 (1995).
29. Cirac, J. I., Zoller, P., Kimble, H. J. & Mabuchi, H. Quantum state transfer and entanglement distribution among distant nodes in a quantum network. *Phys. Rev. Lett.* **78**, 3221–3224 (1997).
30. Moehring, D. L. *et al.* Entanglement of single-atom quantum bits at a distance. *Nature* **449**, 68–71 (2007).
31. Idziaszek, Z., Calarco, T. & Zoller, P. Ion-assisted ground-state cooling of a trapped polar molecule. Preprint at (<http://arxiv.org/abs/1008.1858>) (2010).

Acknowledgements This work was supported by IARPA, DARPA, ONR and the NIST Quantum Information Program. We thank M. Biercuk, A. VanDevender, J. Amini, and R. B. Blakestad for their help in assembling parts of the experiment, and we thank U. Warring and R. Simmonds for comments. This paper, a submission of NIST, is not subject to US copyright.

Author Contributions K.R.B. and C.O. participated in the design of the experiment and built the experimental apparatus. K.R.B. collected data, analysed results and wrote the manuscript. Y.C. fabricated the ion trap chip and collected data. A.C.W. maintained laser systems and collected data. D.L. participated in the design of the experiment, collected data and maintained laser systems. D.J.W. participated in the design and analysis of the experiment. All authors discussed the results and the text of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.R.B. (kenton.brown@nist.gov).

Trapped-ion antennae for the transmission of quantum information

M. Harlander¹, R. Lechner¹, M. Brownnutt¹, R. Blatt^{1,2} & W. Hänsel^{1,2†}

More than 100 years ago, Hertz succeeded in transmitting signals over a few metres to a receiving antenna using an electromagnetic oscillator, thus proving the electromagnetic theory¹ developed by Maxwell. Since this seminal work, technology has developed, and various oscillators are now available at the quantum mechanical level. For quantized electromagnetic oscillations, atoms in cavities can be used to couple electric fields^{2,3}. However, a quantum mechanical link between two mechanical oscillators (such as cantilevers^{4,5} or the vibrational modes of trapped atoms⁶ or ions^{7,8}) has been rarely demonstrated and has been achieved only indirectly. Examples include the mechanical transport of atoms carrying quantum information⁹ or the use of spontaneously emitted photons¹⁰. Here we achieve direct coupling between the motional dipoles of separately trapped ions over a distance of 54 micrometres, using the dipole–dipole interaction as a quantum mechanical transmission line¹¹. This interaction is small between single trapped ions, but the coupling is amplified by using additional trapped ions as antennae. With three ions in each well, the interaction is increased by a factor of seven compared to the single-ion case. This enhancement facilitates bridging of larger distances and relaxes the constraints on the miniaturization of trap electrodes. The system provides a building block for quantum computers and opportunities for coupling different types of quantum systems.

The exchange of quantum information between quantum bits (qubits) at remote sites is a key feature required to render quantum computation truly scalable¹². The dipole–dipole interaction offers a link between separate quantum systems without the need to shuttle particles between sites¹³. The interaction strength depends on the orientation and distance of the dipoles, and is in general given by

$$U_{dd} = \frac{1}{4\pi\epsilon_0} \frac{\mathbf{d}_1 \mathbf{d}_2 - 3(\mathbf{d}_1 \mathbf{e}_r)(\mathbf{d}_2 \mathbf{e}_r)}{r^3} \quad (1)$$

where \mathbf{d}_i are the interacting dipoles, r and \mathbf{e}_r denote respectively the magnitude and direction of their separation, and ϵ_0 is the vacuum permittivity.

Here this interaction is explored using ions or ion strings held in two separate potential wells of a linear segmented ion trap (Fig. 1), where the interacting dipoles are produced by the oscillating charges. As the dipole–dipole interaction decreases rapidly with trap separation, it is advantageous to bring the trapping wells as close together as possible. However, the generation of small inter-well distances requires similarly small distances, d , between the ions and the trap electrodes. This requirement runs counter to the effort to keep a larger ion–electrode separation in order to reduce both ‘anomalous heating’^{14,15} with its d^{-4} scaling and the effects of technical noise of the applied voltages¹⁶. Various routes may be taken to balance these competing requirements. In one approach, the heating rate may be reduced through the use of cryogenic temperatures¹⁷. Here, another approach is taken that uses more ions in the individual traps and enables interaction over larger distances. The additional ions work as ‘antennae’ that increase the motional dipole moment at the respective trapping site.

Given longitudinal alignment of the traps with one particle in each well, the dipole–dipole interaction (equation (1)) is:

$$U_{dd} = -\frac{q_1 q_2}{2\pi\epsilon_0} \frac{\Delta z_1 \Delta z_2}{r^3} \quad (2)$$

$$= -\hbar \frac{\Omega_c}{2} (a_1 + a_1^\dagger)(a_2 + a_2^\dagger) \\ \approx -\hbar \frac{\Omega_c}{2} (a_1 a_2^\dagger + a_1^\dagger a_2) \quad (3)$$

with

$$\Omega_c = \frac{q_1 q_2}{2\pi\epsilon_0 \sqrt{m_1 m_2} \omega_1 \omega_2} \frac{1}{r^3} \quad (4)$$

Here q_i and m_i refer to the charge and mass of the particles, Ω_c is the coupling strength, and $\Delta z_i = \sqrt{\frac{\hbar}{2m_i \omega_i}} (a_i + a_i^\dagger)$ denotes the vibrational amplitude of the motion of ion i ; \hbar is $h/2\pi$ (where h is Planck’s constant). The quantum-mechanical creation and annihilation operators, a_i and a_i^\dagger , act on the individual harmonic oscillators with frequencies ω_i . Rapidly oscillating terms have been neglected in equation (3).

At resonance, that is, for $\omega_1 = \omega_2$, the coupling described by equation (3) leads to a complete exchange of motional states between the two ions after time $T_{\text{swap}} = \pi/\Omega_c$. This is analogous to two coupled

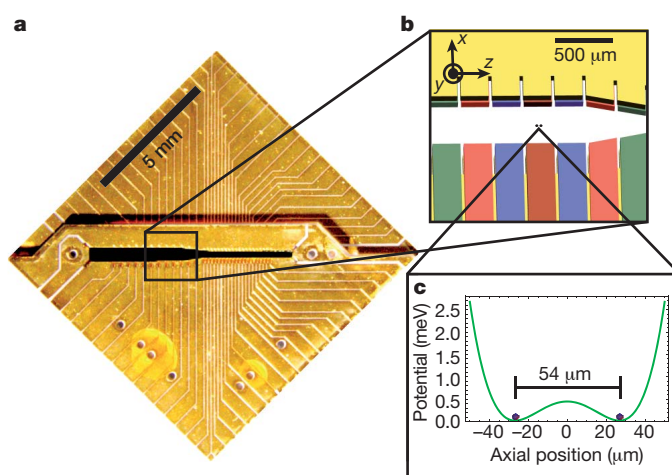


Figure 1 | Creation of trap potential. **a**, Image of the two-layer segmented trap with gold-on-alumina electrodes. The electrodes are 250 μm wide with a pitch of 280 μm , and the ion–electrode spacing is 258 μm . **b**, Schematic close-up of the zone of interest. The electrode pairs used to generate the double-well potential are colour-coded: red for positive voltages, blue for negative voltages and green for compensation voltages. **c**, The axial potential calculated with the parameters determined from the motional spectra in Fig. 4.

¹Institut für Experimentalphysik, Universität Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria. ²Institut für Quantenoptik und Quanteninformation der Österreichischen Akademie der Wissenschaften, Technikerstraße 21a, A-6020 Innsbruck, Austria. [†]Present address: Menlo Systems GmbH, Am Klopferspitz 19a, D-82152, Martinsried, Germany.

pendula connected by a (massless) spring. If one pendulum initially oscillates while the other is at rest, the motion is periodically exchanged between them. The first pendulum comes to a complete stop after a characteristic time, T_{swap} , which, for small coupling, is inversely proportional to the associated spring constant. A quantum-mechanical description may be given using the vibrational quanta, n_i , often labelled 'phonons', in the individual wells. Under the dipole-dipole interaction, an initial motional state $|n_1, n_2\rangle$ becomes entangled with all other motional states $|n'_1, n'_2\rangle$ for which $n'_1 + n'_2 = n_1 + n_2$. Only at odd (even) multiples of T_{swap} is the swapped (original) basis state recovered. Notably, the initial state $|0, 1\rangle$ evolves into the Bell state $(|0, 1\rangle + |1, 0\rangle)/\sqrt{2}$ after time $T_{\text{swap}}/2$, yielding a maximally entangled state of motion. This motional entanglement can be mapped on to the internal electronic state of the ions¹⁸.

In the experiment presented here, the coherent energy exchange is demonstrated between singly charged $^{40}\text{Ca}^+$ ions. They are held in an ion trap with gold-on-alumina electrodes arranged in a two-layer geometry (Fig. 1) similar to the one described in ref. 19. Applying d.c. voltages of up to 110 V to seven adjacent electrode pairs, a double-well potential with a trap separation of 54 μm and an axial trap frequency, $f_0 = 537$ Hz, is created (Methods). The ions are Doppler-cooled on the $S_{1/2}$ - $P_{1/2}$ transition at 397 nm using a single, elliptically shaped laser beam, and detected by collecting the fluorescence light on an electron-multiplied CCD (charge-coupled device) camera and on a photomultiplier tube. Two 729-nm laser beams, individually focused on the two trapping sites, are used to perform sideband cooling on the $S_{1/2}$ - $D_{5/2}$ transition and to map out the sideband spectrum²⁰. During the 4-ms period of sideband cooling, the trapping sites are alternately illuminated with 729-nm light. This alternation is carried out at the approximate rate of the energy exchange to achieve an imbalance of phonon population between the traps. By comparing Rabi-oscillations on the red and blue sideband²¹, the mean phonon number in the first well is calculated. Figure 2 shows the oscillatory behaviour of the phonon number in this well as a function of waiting time. The theoretical fit to the data (Methods) indicates an exchange within $T_{\text{swap}} = 222(10)$ μs and an initial phonon population of $\langle n_2 \rangle = 9(1)$ in the second well; all numbers in parentheses denote 1σ standard deviations with respect to the last digit. Currently, the ions experience a heating rate, $\langle \dot{n} \rangle$, of 1.3(7) quanta per millisecond, which limits the efficiency of the sideband

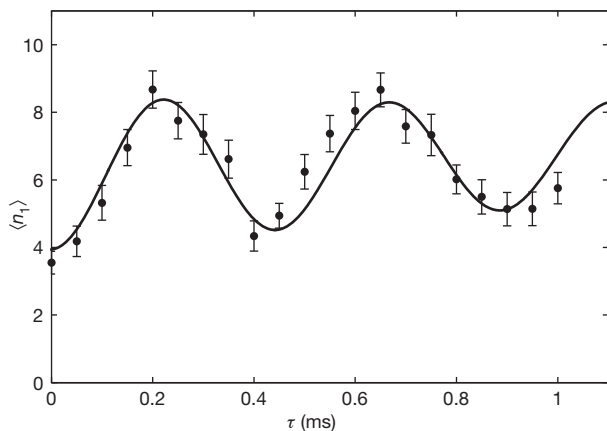


Figure 2 | Energy exchange between two trapped ions over a distance of 54 μm . The data show the average number of phonons in the first trapping well, $\langle n_1 \rangle$, after sideband cooling and a variable waiting time, τ . The sequence for sideband cooling has been arranged to yield different phonon numbers in the two wells, revealing the state exchange as an oscillation of phonon numbers at the level of a few quanta. The observed time for a complete exchange is $T_{\text{swap}} = 222(10)$ μs , indicating a mode splitting of $\Omega_c \approx 2\pi \times 2.25(4)$ kHz. A damping constant $\tau_{\text{damp}} = 3(2)$ ms and a constant background heating, $\langle \dot{n} \rangle$, of 1.3(7) quanta per millisecond are inferred from the fit to the data. The error bars indicate one standard deviation as inferred from Monte Carlo simulations. Lateral deviations of the data from the fit are attributed to small drifts in the resonance condition which modify the exchange rate.

cooling and is comparable to other room-temperature traps of similar size²².

Ideally, the exchange rate would be significantly larger than the average heating rate. To enhance the exchange rate, strings of several ions can be used in the trapping wells. Equation (4) suggests that the coupling strength scales in proportion to the number of ions. This assumes the ion strings to be point-like objects. In practice, the extended nature of the strings can cause a significantly faster increase in the coupling strength.

This increase is demonstrated by mapping out the ions' excitation spectrum as the trapping frequencies of the individual sites are scanned through the resonance condition. The frequency scan is achieved by applying a control voltage U_{ax} to an outer trap-electrode pair (Methods). The dipole-dipole coupling manifests itself as an avoided crossing, separating the mode frequencies by $\Omega_c/2\pi$. Close to resonance, the motion of the ion strings is strongly coupled, and the oscillation can be excited with 729-nm light on either of the two trapping sites²⁰. Figure 3a shows an example of this avoided crossing measured with two ions in each well, while Fig. 3b represents an individual sideband spectrum. The motional spectra of five ion configurations have been analysed, using up to three ions in each well. The configurations and corresponding mode spectra are displayed in Fig. 4. The curves present calculations using a common set of fit parameters for the potential and for the action of the control voltage U_{ax} (Methods), and explicitly include the extended nature of the ion strings. For the configuration with one ion per well, the observed splitting of 1.9 kHz agrees within one standard deviation with the energy-exchange rate from Fig. 2, which was measured with the same trap parameters. The data further show that the splitting is increased from 1.9(3) kHz to 14(1) kHz by using up to three ions in each well, without any modification of the external potential. The sevenfold increase of the coupling is beyond the factor of three that is expected from a simple point-charge model and is due to the anharmonicity of the individual potential wells: as the outer potential walls are steeper than the inner ones, the ion-strings' centres of mass get closer as more ions are added. At the same time, the average oscillation frequency is reduced. Both of these effects lead to an increase in the coupling strength (equation (4)). The extent of the ion strings provides a comparable increase, as it is non-negligible with respect to the inter-well distance.

The demonstrated coupling can be used in diverse schemes to create entanglement or to perform gates. (See also independent experiments at NIST Boulder with trapped $^9\text{Be}^+$ ions²³.) The creation of Bell states

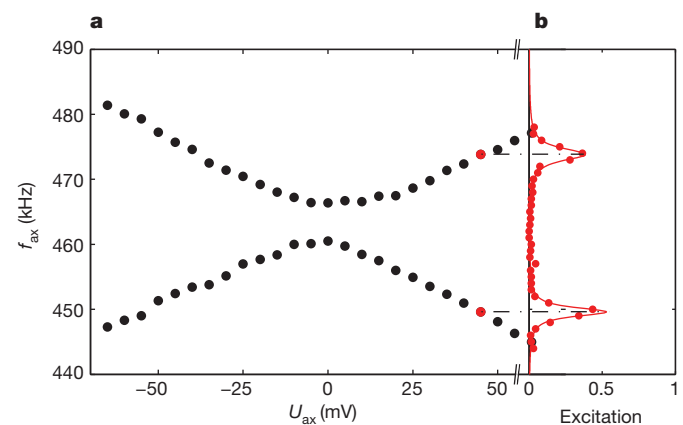


Figure 3 | Demonstration of avoided crossing derived from sideband spectra. a, Oscillation frequencies of four trapped ions (two in each well) as a function of the axial control voltage U_{ax} , yielding a mode splitting of 5.5(3) kHz. The data points (filled circles) correspond to peaks in individual sideband spectra taken on the $S_{1/2}$ - $D_{5/2}$ transition. The error bars are smaller than the filled circles. b, Example of an individual sideband spectrum. The resonant frequencies are marked with a dot-dashed line, and the corresponding data points in a are coloured red.

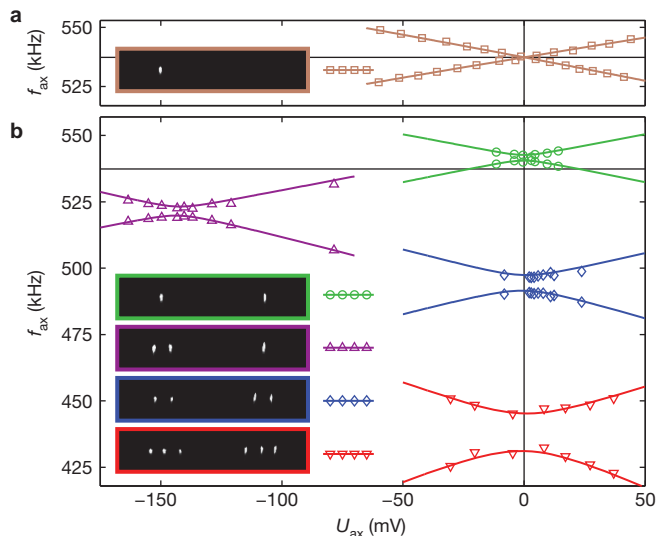


Figure 4 | Experimentally observed dipole-dipole coupling for various ion configurations in a double-well potential. The graphs display the oscillation frequencies of the two lowest vibrational modes as a function of control voltage U_{ax} and reveal the mode splitting at resonance; the horizontal line in each panel shows f_0 . **a**, Unmodified trap frequencies, as measured with only a single ion in the left well (negative slope) or in the right well (positive slope). **b**, Spectra with up to three ions in each well, as depicted in the insets (inset border colour matches data symbol and line colour). The lines represent predictions from numerical calculations and fit all data simultaneously. A small drift in the control voltage of 7 mV over one hour has been taken into account, aligning the spectra on top of one another. The mode splittings (kHz) for the four different configurations are 1.9(3), 2.4(9), 6.0(3) and 14(1). The shift to lower voltages for three ions is due to the asymmetry in their configuration, (2+1). The vertical elongation of the ion pictures is due to aberrations arising from an off-axis position of the ions relative to the imaging axis.

is discussed above and requires ground-state cooling of the ions. One route to perform a thermally robust quantum gate has been described¹¹, but it requires a state-dependent pushing force by an appropriate laser. It should be noted that a fast variant²⁴ of the thermally robust Mølmer-Sørensen gate²⁵ may also be implemented in the system under consideration: if the required bichromatic laser beams are tuned to the centre of the mode splitting, a gate can be performed within $T_{\text{gate}} = 4\pi/\Omega_c$.

Our coupling scheme also has implications for possible architectures for quantum computation. It is straightforward to imagine a linear array of ion traps, each holding a string of ions, able to interact axially. One could further imagine a two-dimensional array of traps in which parallel dipoles can also be coupled. Bringing rows of traps pairwise into resonance allows the creation of linear cluster states within two steps. A two-dimensional cluster state would only require two further steps by pair-wise coupling of columns²⁶.

The amplification of the dipole-dipole interaction is not limited to ions; the presented technique may be directly transferred to the coupling of trapped Rydberg atoms²⁷. Nor is the interaction limited to systems of equal mass, as demonstrated by the mode-splitting for the asymmetric configuration (2+1) (Fig. 4). Atomic and ionic systems may thus be combined. Neutral atoms have already been brought close to ions^{28,29}. When the ions' oscillation frequency is tuned into resonance with adjacent levels of a Rydberg atom, coupling between ions and Rydberg atoms may become achievable.

METHODS SUMMARY

Calculation of eigenfrequencies in the double-well potential. A symmetric double-well potential is created along the trap axis by applying 2.8 V, 110.4 V, −16.8 V, 13.7 V, −16.8 V, 110.4 V and −33.0 V to the seven adjacent electrode pairs marked in Fig. 1b. The control voltage, U_{ax} , is additionally applied to the leftmost electrode pair, creating a nearly homogeneous electric field in the centre

region. This shifts the trap frequencies of the two wells in opposite directions and is used to tune the wells through resonance. In the region of interest the double-well potential is well described by a fourth-order polynomial, where the zeroth order term does not influence the dynamics and the third-order coefficient is eliminated by appropriate choice of origin:

$$E_{\text{pot}}(z) = q(\alpha_1 z + \alpha_2 z^2 + \alpha_4 z^4) \quad (5)$$

The polynomial coefficients, α_2 and α_4 , determine both the uncoupled trap frequency, ω_0 , and the inter-well separation, r , at the symmetry point³⁰ ($\alpha_1 = 0$). The action of the tuning voltage is modelled as a linear contribution to the coefficients, whereby $\alpha_1 = 4.61 \times 10^{-24} \text{ m}^{-1} \times U_{ax}$ and $\alpha_2 = \alpha_2(U_{ax} = 0) + 0.9 \times 10^{-20} \text{ m}^{-2} \times U_{ax}$. The equilibrium positions for N ions, $\mathbf{z}_0^N = (z_{0,1}, \dots, z_{0,N})$, are those minimizing the total energy, E_{tot}^N :

$$E_{\text{tot}}^N(\mathbf{z}^N) = \sum_{i=1}^N E_{\text{pot}}(z_i) + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{q^2}{4\pi\epsilon_0} \times \frac{1}{|z_i - z_j|} \quad (6)$$

The theoretical data for the vibrational mode frequencies, ω_b , in Fig. 4 are computed from the two lowest eigenvalues of the Hessian matrix of $E_{\text{tot}}^N(\mathbf{z}^N)$ at positions \mathbf{z}_0^N .

Energy exchange. The theoretical fit to the data in Fig. 2 is calculated using

$$\langle n_1 \rangle(\tau) = \langle n_1 \rangle(0) + \frac{1}{2} \langle \Delta n \rangle \left(1 - \cos\left(\pi \frac{\tau}{T_{\text{swap}}}\right) \right) \exp\left(-\frac{\tau}{\tau_{\text{damp}}}\right) + \langle \dot{n} \rangle \tau \quad (7)$$

with $\langle n_1 \rangle(0)$ denoting the initial mean phonon number in the first well, and $\langle \Delta n \rangle = \langle n_2 \rangle(0) - \langle n_1 \rangle(0)$ representing the difference between the initial mean phonon numbers in each well. T_{swap} is the time required for a complete exchange of motional states. Damping due to decoherence and any slow fluctuations of the experimental parameters is taken into account with τ_{damp} , and $\langle \dot{n} \rangle$ denotes the background heating rate.

Received 12 November 2010; accepted 4 January 2011.

Published online 23 February 2011.

- Hertz, H. Über die Grundgleichungen der Electrodynamik für bewegte Körper. *Ann. Phys. Chem.* **277**, 369–399 (1890).
- Kuhr, S. et al. Ultrahigh finesse Fabry-Perot superconducting resonator. *Appl. Phys. Lett.* **90**, 164101 (2007).
- Gleyzes, S. et al. Quantum jumps of light recording the birth and death of a photon in a cavity. *Nature* **446**, 297–300 (2007).
- Anetsberger, G. et al. Near-field cavity optomechanics with nanomechanical oscillators. *Nature Phys.* **5**, 909–914 (2009).
- Schliesser, A., Del'Haye, P., Nooshi, N., Vahala, K. J. & Kippenberg, T. J. Radiation pressure cooling of a micromechanical oscillator using dynamical backaction. *Phys. Rev. Lett.* **97**, 243905 (2006).
- Kinoshita, T., Wenger, T. & Weiss, D. S. A quantum Newton's cradle. *Nature* **440**, 900–903 (2006).
- Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75**, 281–324 (2003).
- Monroe, C., Meekhof, D. M., King, B. E. & Wineland, D. J. A “Schrödinger Cat” superposition state of an atom. *Science* **272**, 1131–1136 (1996).
- Jost, J. D. et al. Entangled mechanical oscillators. *Nature* **459**, 683–685 (2009).
- Blinov, B. B., Moehring, D. L., Duan, L.-M. & Monroe, C. Observation of entanglement between a single trapped atom and a single photon. *Nature* **428**, 153–157 (2004).
- Cirac, J. I. & Zoller, P. A scalable quantum computer with ions in an array of microtraps. *Nature* **404**, 579–581 (2000).
- DiVincenzo, D. P. The physical implementation of quantum computation. *Fortschr. Phys.* **48**, 771–783 (2000).
- Kielinski, D., Monroe, C. & Wineland, D. J. Architecture for a large-scale ion-trap quantum computer. *Nature* **417**, 709–711 (2002).
- Deslauriers, L. et al. Scaling and suppression of anomalous heating in ion traps. *Phys. Rev. Lett.* **97**, 103007 (2006).
- Seidelin, S. et al. Microfabricated surface-electrode ion trap for scalable quantum information processing. *Phys. Rev. Lett.* **96**, 253003 (2006).
- Leibbrandt, D. R. et al. Modeling ion trap thermal noise coherence. *Quant. Inform. Comput.* **7**, 052–072 (2007).
- Labaziewicz, J. et al. Suppression of heating rates in cryogenic surface-electrode ion traps. *Phys. Rev. Lett.* **100**, 013001 (2008).
- Cirac, J. I. & Zoller, P. Quantum computations with cold trapped ions. *Phys. Rev. Lett.* **74**, 4091–4094 (1995).
- Schulz, S. A., Poschinger, U., Ziesel, F. & Schmidt-Kaler, F. Sideband cooling and coherent dynamics in a microchip multi-segmented ion trap. *N. J. Phys.* **10**, 045007 (2008).
- Roos, C. F. et al. Quantum state engineering on an optical transition and decoherence in a Paul trap. *Phys. Rev. Lett.* **83**, 4713–4716 (1999).
- Meekhof, D. M., Monroe, C., King, B. E., Itano, W. M. & Wineland, D. J. Generation of nonclassical motional states of a trapped atom. *Phys. Rev. Lett.* **76**, 1796–1799 (1996).
- Turchette, Q. A. et al. Heating of trapped ions from the quantum ground state. *Phys. Rev. A* **61**, 063418 (2000).
- Brown, K. R. et al. Coupled quantized mechanical oscillators. *Nature* doi:10.1038/nature09721 (this issue).

24. Benhelm, J. *et al.* Towards fault-tolerant quantum computing with trapped ions. *Nature Phys.* **4**, 463–466 (2008).
25. Sørensen, A. & Mølmer, K. Quantum computation with ions in thermal motion. *Phys. Rev. Lett.* **82**, 1971–1974 (1999).
26. Raussendorf, R., Browne, D. E. & Briegel, H. J. Measurement-based quantum computation on cluster states. *Phys. Rev. A* **68**, 022312 (2003).
27. Urban, E. *et al.* Observation of Rydberg blockade between two atoms. *Nature Phys.* **5**, 110–114 (2009).
28. Zipkes, C., Palzer, S., Sias, C. & Köhl, M. A trapped single ion inside a Bose-Einstein condensate. *Nature* **464**, 388–391 (2010).
29. Schmid, S., Härter, A. & Denschlag, J. H. Dynamics of a cold trapped ion in a Bose-Einstein condensate. *Phys. Rev. Lett.* **105**, 133202 (2010).
30. Home, J. P. & Steane, A. M. Electrode configuration for fast separation of trapped ions. *Quant. Inform. Comput.* **6**, 289–325 (2006).

Acknowledgements We thank H. Häffner for discussions at an early state of the project. We acknowledge the support of the EU STREP project MICROTRAP, the Austrian Science Fund (FWF), the EU network SCALA, the European Research Council (ERC) and the Institut für Quanteninformation GmbH.

Author Contributions The experiments were performed by M.H., R.L. and W.H.; M.H., M.B., R.L., W.H. and R.B. contributed to the set-up; the data analysis was performed by M.H. and W.H.; the original idea was devised by W.H. and R.B.; and all authors contributed to the discussion of the results and participated in manuscript preparation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.B. (rainer.blatt@uibk.ac.at).

Circuit cavity electromechanics in the strong-coupling regime

J. D. Teufel¹, Dale Li¹, M. S. Allman¹, K. Cicak¹, A. J. Sirois¹, J. D. Whittaker¹ & R. W. Simmonds¹

Demonstrating and exploiting the quantum nature of macroscopic mechanical objects would help us to investigate directly the limitations of quantum-based measurements and quantum information protocols, as well as to test long-standing questions about macroscopic quantum coherence^{1–3}. Central to this effort is the necessity of long-lived mechanical states. Previous efforts have witnessed quantum behaviour⁴, but for a low-quality-factor mechanical system. The field of cavity optomechanics and electromechanics^{5,6}, in which a high-quality-factor mechanical oscillator is parametrically coupled to an electromagnetic cavity resonance, provides a practical architecture for cooling, manipulation and detection of motion at the quantum level¹. One requirement is strong coupling^{7–9}, in which the interaction between the two systems is faster than the dissipation of energy from either system. Here, by incorporating a free-standing, flexible aluminium membrane into a lumped-element superconducting resonant cavity, we have increased the single-photon coupling strength between these two systems by more than two orders of magnitude, compared to previously obtained coupling strengths. A parametric drive tone at the difference frequency between the mechanical oscillator and the cavity resonance dramatically increases the overall coupling strength, allowing us to completely enter the quantum-enabled, strong-coupling regime. This is evidenced by a maximum normal-mode splitting of nearly six bare cavity linewidths. Spectroscopic measurements of these ‘dressed states’ are in excellent quantitative agreement with recent theoretical predictions^{10,11}. The basic circuit architecture presented here provides a feasible path to ground-state cooling and subsequent coherent control and measurement of long-lived quantum states of mechanical motion.

Despite the prevalence of mechanical oscillators in many practical technological applications, demonstrating and using robust quantum behaviour in these systems has proven exceedingly difficult. Recent experimental and theoretical progress has shown that mechanical states may outlive those produced in other media, making them useful for quantum information processing or for macroscopic tests of the quantum–classical divide. Remarkably, a 6 GHz piezoelectric dilatation resonator with strong capacitive coupling to a superconducting phase qubit, cooled by a dilution refrigerator, has allowed control and measurement of a single microwave phonon⁴. This important result confirms the ability to manipulate mechanical quantum states, but the phonon states were short-lived, surviving for only ~ 6 ns. For mechanical systems to be practically useful, they must survive much longer than a controlled operation time, which in this case is also about ~ 6 ns. It is a known difficulty that mechanical resonator performance degrades considerably as the fundamental mode frequency increases¹² and that piezoelectric materials can suffer from high internal losses⁴.

These difficulties can be alleviated by using low-frequency (< 100 MHz), free-standing flexural oscillators, but they must also be cooled further, below 5 mK, in order to avoid thermal decoherence. This requires active cooling. An analogous situation is encountered for charged ions and their motion in a trapping potential; here, a sideband cooling technique brings the system to its motional ground state¹³. The interaction between different ions through their long-lived collective

motion is crucial to facilitate quantum information processing. Optomechanics strives to achieve this type of performance, but with ion motion replaced by a macroscopically large mechanical oscillator, and the internal states of ions replaced by a high-frequency, high-finesse optical Fabry–Pérot cavity. The main difficulty with this approach has been the lack of strong coupling between the two disparate systems, something easily achieved with trapped ions. Recently, strong coupling was narrowly achieved with an optomechanical system¹⁴ under large optical driving. Despite the strong coupling, this system was not in the quantum-enabled regime, as the thermal decoherence rate Γ_{th} was too large for observing quantum effects ($\Gamma_{\text{th}} = n_m \Gamma_m \gg \kappa$, where κ and Γ_m are the intrinsic dissipation rates of the cavity and the mechanical oscillator, respectively). In this work, we describe a macroscopic, electromechanical system compatible with cryogenic operation that achieves a coupling rate g that is much greater than κ and Γ_m . As this device operates in both quantum-enabled and strong-coupling regimes ($\Gamma_{\text{th}} \ll \kappa < g < \Omega_m$, where Ω_m is the angular resonance frequency of the mechanical mode), it simultaneously fulfils all the requirements for observing and controlling many theoretically predicted quantum effects^{2,3,7–9}.

Cavity electromechanical systems operate under the same principles as their cavity optomechanical analogues. The motion of the mechanical oscillator modulates the frequency ω_c of the optical or electromagnetic cavity. The single-photon coupling strength $g_0 = Gx_{\text{zp}}$ is the product of $G = d\omega_c/dx$ (the change in the cavity frequency for a given displacement x) and the zero point motion for mass m , given by $x_{\text{zp}} = \sqrt{\hbar/2m\Omega_m}$ (\hbar is $h/2\pi$, where h is Planck’s constant). In order to maximize the coupling, we would like to maximize the radiation pressure force per photon ($\hbar G$) and use a light mechanical object with a low resonant mode frequency Ω_m (a soft spring constant). Optomechanical systems with silica microtoroids¹¹, dielectric membranes¹⁵ and deflecting micromirrors¹⁴ have all attempted to balance these preferred characteristics. The move to an electromechanical system has some advantages. These systems utilize low-loss superconducting circuits that are easily cooled to temperatures below 100 mK. Pioneering microwave electromechanical systems using parametric transducers achieved relatively large values of G and continue their development towards detection of gravitational waves^{16,17}. Recent progress with lithographically fabricated microwave resonators¹⁸ has focused on using nanowires to achieve large x_{zp} , and has also enabled sideband cooling^{19,20} and near quantum-limited detection^{21,22}, despite relatively weak coupling. Our approach, described below, takes advantage of the large capacitance produced by suspending an aluminium membrane approximately 50 nm above an aluminium base electrode (Fig. 1), to optimize both G and x_{zp} to achieve the largest g_0 to date.

Imagine two parallel metal plates of area A separated by a uniform vacuum gap d and electrically connected with a coiled wire (Fig. 1a, b). These plates form a capacitance $C = A\epsilon_0/d$ (where ϵ_0 is the permittivity of free space) that electrically resonates with the inductance L of the coil at an angular frequency $\omega_c = 1/\sqrt{LC}$. When the top plate is free to move a small distance $x \ll d$, then $G = d\omega_c/dx \approx -\omega_c/(2d)$. For a cavity with $\omega_c/2\pi = 10$ GHz and a plate separation of 50 nm,

¹National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA.

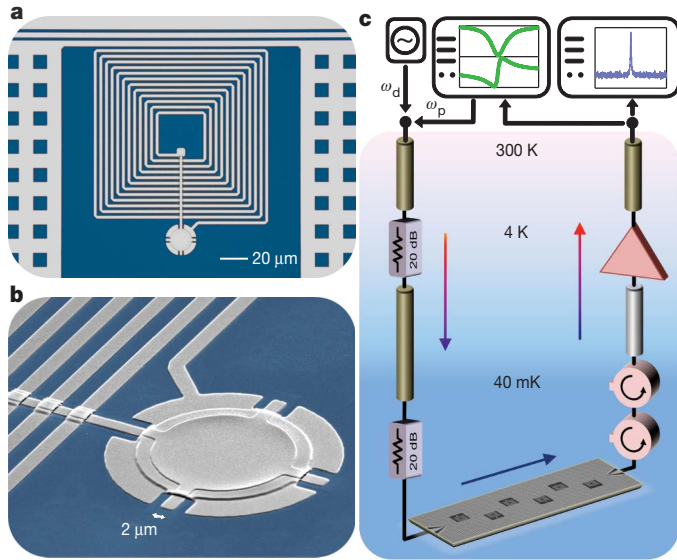


Figure 1 | Schematic description of the experiment. **a**, False-colour optical micrograph of the microwave resonator formed by a spiral inductor shunted by a parallel-plate capacitor. A coplanar waveguide transmission line (top) inductively couples microwave signals to and from the resonator (centre). **b**, False-colour scanning electron micrograph showing the upper plate of the capacitor suspended ~ 50 nm above the lower plate and free to vibrate like a taut, circular drum. The metallization is sputtered aluminium (grey) patterned on a sapphire substrate (blue). **c**, This circuit is measured by applying a microwave tone near the electrical resonance frequency through coaxial lines. Cryogenic attenuators (20 dB) on the input line and isolators (circular structures) on the output line ensure that thermal noise is reduced below the vacuum noise at microwave frequencies. The measured signal encoding the mechanical motion as modulation sidebands is transmitted to a low noise, cryogenic amplifier (triangle) via a superconducting coaxial cable and demodulated at room temperature with either a vector network analyser (green trace) or a spectrum analyser (blue trace).

$G/2\pi \approx 100 \text{ MHz nm}^{-1}$. This is to be contrasted with previous experiments that used nanowires of very low mass and high aspect ratio^{18–22}. These wires only contribute a fraction η ($< 1/1,000$) of the total capacitance so that $G \approx -\eta\omega_c/(2d)$. So although these wires have large zero-point motion, the sensitivity is limited to $G/2\pi < 100 \text{ kHz nm}^{-1}$.

Our circuit is fabricated with wafer-scale optical lithographic techniques developed for creating low-loss vacuum-gap-based microwave components²³. The nearly circular membrane is 100 nm thick and has a diameter of 15 μm , allowing drum-like modes to resonate freely. The fundamental mode is $\Omega_m/2\pi = 10.69 \text{ MHz}$, giving a zero-point motion of $x_{zp} = 4.1 \text{ fm}$. The total capacitance $C \approx 38 \text{ fF}$ combined with a parallel inductance, $L \approx 12 \text{ nH}$, provides a fundamental microwave cavity resonance²³ of $\omega_c/2\pi \approx 7.5 \text{ GHz}$. The device is cooled to 40 mK, far below the superconducting transition temperature of aluminium. To measure the motion of the membrane, we apply microwave signals through heavily attenuated coaxial lines, which inductively couple to the superconducting cavity, as shown schematically in Fig. 1c. The transmitted signals are amplified at 4 K with a cryogenic low-noise amplifier and demodulated at room temperature with either a commercial vector network analyser (for characterizing the cavity mode) or a spectrum analyser (for characterizing the mechanical mode).

Figure 2a shows the magnitude of transmission, $|T|^2$, near the cavity resonance at sufficiently low microwave power that radiation pressure effects can be neglected. A Lorentzian fit yields a resonance frequency of $\omega_c/2\pi = 7.47 \text{ GHz}$ and a loaded intensity decay rate of $\kappa/2\pi = 170 \text{ kHz}$. The depth of the dip at resonance shows that the circuit is overcoupled, so that the dominant source of damping is the intentional inductive coupling to the transmission line, $\kappa_{\text{ex}}/2\pi = 130 \text{ kHz}$, which is much greater than the intrinsic decay rate,

$\kappa_0/2\pi = 40 \text{ kHz}$. The motion of the drum mode modulates the capacitance and thus the frequency of the electrical resonator, creating sidebands above and below the microwave drive frequency at $\omega_d \pm \Omega_m$. Figure 2b shows the noise power of the upper sideband due to the thermal motion of the drum at its fundamental mode, $\Omega_m/2\pi = 10.69 \text{ MHz}$. These data show that the mechanical resonance has an intrinsic damping rate of $\Gamma_m/2\pi = 30 \text{ Hz}$ and a high mechanical quality factor of $Q_m = 360,000$. Although the microscopic mechanisms responsible for mechanical dissipation need more investigation, these values are consistent with the improved Q -values associated with tensile stress found with aluminium nanowire-based oscillators^{18,19,21}. With $\Omega_m/\kappa = 63$, it is clear that we are deeply within the resolved-sideband regime where the mechanical resonance frequency is much larger than the cavity linewidth: this is a prerequisite for both sideband cooling to the ground state and for observing normal-mode splitting^{7–9}.

The quantum-mechanical behaviour of this parametrically coupled system is described by the interaction Hamiltonian, $H_I = -\hbar a^\dagger a g_0 (b^\dagger + b)$, where a^\dagger (a) and b^\dagger (b) are the creation (annihilation) operators for photons and phonons, respectively. Because the motion of this drum strongly influences ω_c , microwave signals can be used not only to detect the motion of the oscillator but also to impart back-action forces on it. The radiation pressure force of the microwave drive photons gives rise to ‘optical’ spring and damping effects. The interaction Hamiltonian can be linearized in a frame co-rotating with the drive, taking the form, $H_I = -\hbar g (a^\dagger + a)(b^\dagger + b)$, where $g = g_0 \sqrt{n_d}$ is the total (linearized optomechanical) coupling strength^{7–9}, and n_d is

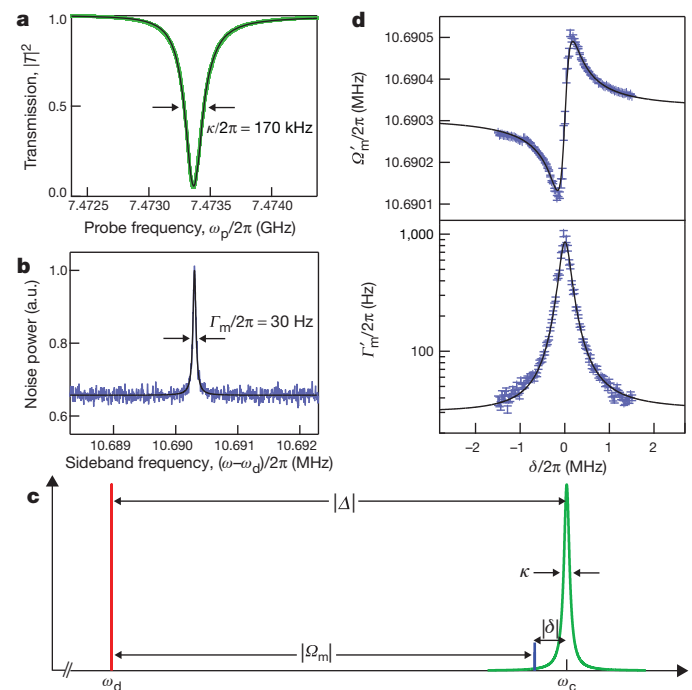


Figure 2 | Characterization of mechanical and microwave resonances. **a**, Measured probe transmission spectrum (green) and Lorentzian fit (black) of the microwave circuit at low power, where optomechanical effects are negligible. The width of the resonance yields the overcoupled, intensity decay rate $\kappa/2\pi = 170 \text{ kHz}$. **b**, The mechanical resonance manifests itself as a peak in the noise spectrum (blue), which appears Ω_m above and below the microwave drive frequency, owing to the thermal motion of the drum up- or down-converting microwave photons. At low microwave power, where back-action effects are negligible, the Lorentzian fit (black) yields an intrinsic mechanical dissipation rate $\Gamma_m/2\pi = 30 \text{ Hz}$ ($Q_m = 360,000$). **c**, Schematic diagram of the relative frequencies of the microwave drive (red) and the upper mechanical sideband (blue) with respect to the narrow cavity resonance (green). **d**, The modified mechanical resonance frequency Ω'_m and damping rate Γ'_m as a function of the relative detuning δ fit well to the theory of dynamical back-action (black), yielding $G/2\pi = 56 \pm 7 \text{ MHz nm}^{-1}$.

the number of drive photons in the cavity. If the drive is detuned so that its upper mechanical sideband is near the cavity resonance, $\delta = (\omega_d + \Omega_m) - \omega_c \ll \Omega_m$ (Fig. 2c), the modified mechanical resonance frequency Ω'_m and damping rate Γ'_m closely follow the imaginary and real parts of the cavity response. In the resolved sideband regime these quantities are well approximated by^{7,8,24}:

$$\Omega'_m \approx \Omega_m + \frac{4g^2\delta}{\kappa^2 + 4\delta^2} \quad (1)$$

$$\Gamma'_m \approx \Gamma_m + \frac{4g^2\kappa}{\kappa^2 + 4\delta^2} \quad (2)$$

Figure 2d shows the measured effects of this dynamical back-action on the drum as a function of δ . The incident microwave power P_{in} is held constant at 10 pW. As this power is applied very far from the cavity resonance, it results in a greatly reduced number of photons in the cavity, given by $n_d = (2P_{\text{in}}\kappa_{\text{ex}}/\hbar\omega_d)/(\kappa^2 + 4\Delta^2)$, where $\Delta = \omega_d - \omega_c$. Even for this moderate-power microwave drive with $n_d \leq 800$, the effects on the mechanical oscillator are quite striking; the intrinsic mechanical damping is dominated by the damping from the microwave photons. Knowing n_d and estimating x_{zp} from geometry, we can fit these data to equations (1) and (2) (shown in black in Fig. 2d) to extract $G/2\pi = 56 \pm 7 \text{ MHz nm}^{-1}$.

Just as the microwave photons strongly affect the mechanical mode, the symmetry of the parametric interaction suggests that the mechanics should influence the cavity mode. To investigate this, we apply both a microwave drive tone ω_d and a second probe tone ω_p . Here, the drive tone will induce an interaction between the mechanics and the cavity, while the probe tone will monitor the response of the cavity. This technique provides a way to measure the spectroscopy of the ‘dressed’ cavity states in the presence of the electromechanical interaction. Figure 3a shows a series of cavity probe spectra taken with successively higher microwave power applied with $\Delta = -\Omega_m$. Once the drive power is large enough that $g \approx \sqrt{\Gamma_m\kappa}$, the mechanical sideband of the driving field appears in the cavity response function. As n_d , and hence g , increase, so does the normalized probe transmission at the cavity resonance, $|T(\omega_c)|$. The width of this peak also increases and is given by the modified mechanical damping rate in equation (2). This electromechanical effect can be understood as the result of a radiation pressure force at the beat frequency between the drive and probe photons, which drives the motion of the drum near its resonance frequency. The driven motion results in a mechanical sideband on the drive field that can interfere with the probe field and hence modifies the probe spectrum. This interference is the mechanical analogue of electromagnetically induced transparency²⁵, which is well known in atomic physics and has enabled such innovations as slow light and photon storage. Recently, these effects have been addressed in the context of optomechanics, both theoretically^{10,11} and experimentally¹¹. The analogous effect is demonstrated here for the first time in an electromechanical system. For our electrical circuit, the transmission spectrum is^{10,11}:

$$T = 1 - \frac{\kappa_{\text{ex}}(1 - j\chi)}{\kappa + 2j(\omega_p - \omega_c) + 4\chi(\omega_d - \omega_c)} \quad (3)$$

where

$$\chi = \frac{4g^2\Omega_m}{[\kappa + 2j(\omega_p - 2\omega_d + \omega_c)][\Omega_m^2 - (\omega_p - \omega_d)^2 + j(\omega_p - \omega_d)\Gamma_m]}$$

At high enough power, Γ'_m becomes comparable to or greater than κ , at which point equations (1) and (2) are no longer valid. This is precisely the point at which the driven system enters the strong-coupling regime, in which the coupling exceeds the intrinsic dissipation in either of the original modes ($g \geq \kappa \gg \Gamma_m$). The eigenmodes of the driven, coupled system are now hybrids of the original radio-frequency mechanical resonance and the microwave electrical resonance. The

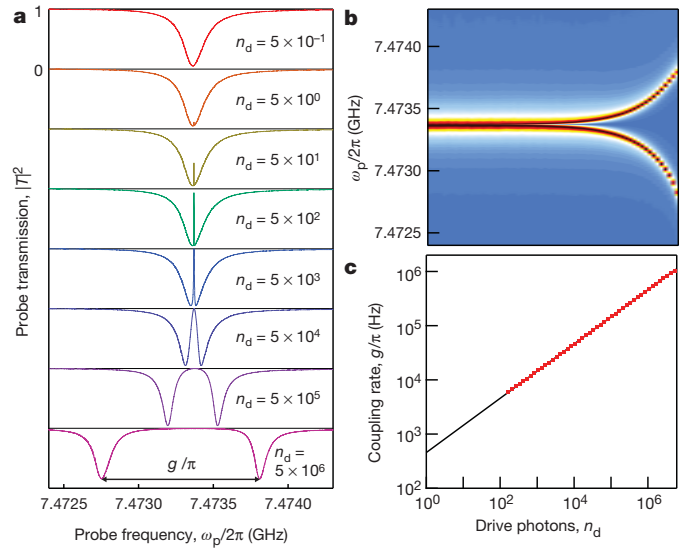


Figure 3 | Demonstration of the strong-coupling regime. **a**, Normalized microwave cavity transmission in the presence of a microwave drive applied with $\Delta = -\Omega_m$, with successive plots for increasing drive amplitude n_d . At moderate drive amplitude ($n_d \approx 10$), the interference between the drive and probe photons results in a narrow peak in the cavity spectrum, whose width is given by the mechanical linewidth Γ'_m . This represents the onset of electromechanically induced transparency¹¹. When Γ'_m becomes comparable to κ , the cavity resonance splits into normal modes. The eigenmodes of the system are no longer purely electrical or mechanical, but are a pair of hybrid electromechanical resonances. **b**, The transmission as a function of probe frequency and number of drive photons shows that the driven system enters the strong-coupling regime ($g \geq \kappa, \Gamma_m$). Here, the logarithmic colour scale shows the transmission as it varies from -13 dB (dark red) to 0 dB (blue). **c**, The measured coupling rate g (red) follows the expected dependence on the number of drive photons, with a fit to $\sqrt{n_d}$ shown in black.

system exhibits the well-known normal-mode splitting of two strongly coupled harmonic oscillators. This effect is observed through the cavity response instead of the mechanical response as is done in other experiments¹⁴. For our device, progression into the strong-coupling regime is shown in Fig. 3a, b with a crossover occurring at $n_d \approx 10^5$. In this regime, the damping rate of each of the two normal modes approaches $(\kappa + \Gamma_m)/2$, and the coupling results in a splitting of $2g$. In Fig. 3c, g is extracted by fitting each spectrum at a given drive power to equation (3). The splitting shows $\sqrt{n_d}$ dependence, with a single-photon coupling rate of $g_0/\pi = 460 \text{ Hz}$. At the highest drive power, $n_d = 5 \times 10^6$, the splitting is $g/\pi = 1 \text{ MHz}$, exceeding both κ and Γ_m . The cooperativity $\mathcal{C} = 4g^2/\Gamma_m\kappa$ represents a good figure of merit for an opto- or electromechanical system, regardless of its detailed construction. For our system, we find a maximum cooperativity of $\mathcal{C} \approx 200,000$, beyond those previously achieved^{11,14,15,20,21}.

By measuring the in-phase and quadrature components of the transmitted probe signal, we determine the real and imaginary parts of the cavity spectrum in the presence of the electromechanical interactions. We find excellent agreement between theory and experiment for both magnitude and phase over the entire range of accessible detunings and powers. Figure 4a shows the magnitude and phase of the probe transmission for three arbitrary values of δ when $g/\pi = 150 \text{ kHz}$. The black lines are fits to the magnitude and argument of equation (3). The narrow resonance resulting from the mechanical interaction could be used as a tunable delay for microwave signals¹¹, with a maximum group delay of $\sim 5 \text{ ms}$. Two-dimensional plots of $|T|$ are shown in Fig. 4 as a function of both ω_d and ω_p . At low drive amplitude (Fig. 4b), the narrow mechanical sideband appears as a sharp dip or peak in transmission whenever $\omega_p = \omega_d + \Omega'_m$. As the drive amplitude is increased (Fig. 4c), the dispersive coupling between the normal modes becomes apparent. Although the mechanical sideband is narrow when it is far

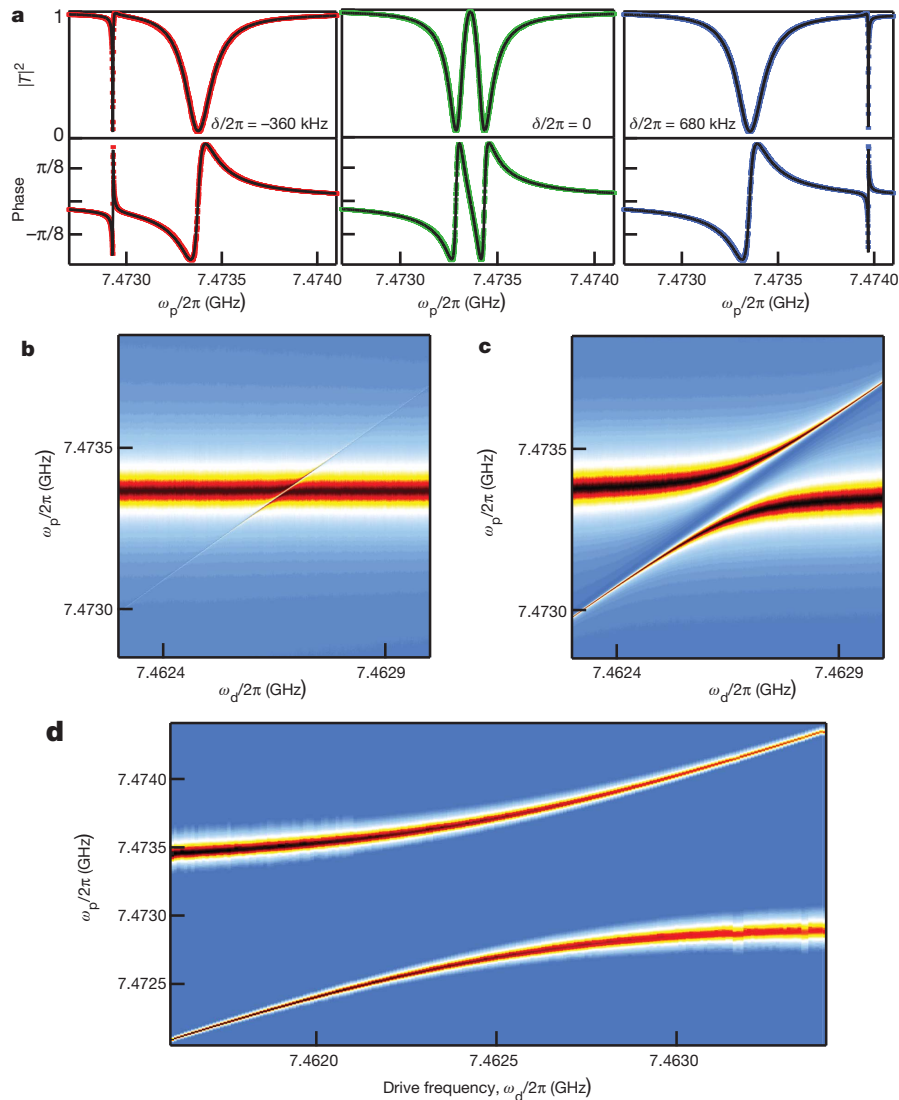


Figure 4 | Spectroscopy in the strong-coupled regime. **a**, The normalized magnitude and phase of the cavity transmission in the presence of a strong microwave drive. The data shown in red, green and blue are for three different detunings δ with fits to equation (3) (black). **b**, For relatively low drive power ($n_d \approx 10$), the mechanical sideband appears as a sharp dip or peak in the probe

from resonance, it inherits the cavity's larger damping rate as it approaches ω_c . At the largest coupling, the normal-mode splitting is appreciable for a broad range of drive and probe frequencies (Fig. 4d).

This experiment demonstrates that by engineering a microwave resonant circuit with a free-standing micromechanical membrane or drum, a large coupling is possible without degradation of either the mechanical or electromagnetic quality factors. Our measurements have shown clear quantitative agreement with theoretical predictions for dynamical back-action and normal-mode splitting, with the latter being exceptionally large, namely 10% of Ω_m . Simple refinements—such as reducing d , using thinner membranes or exploring higher-order modes—could further enhance the coupling strength and relax cooling requirements. This new system shows the performance necessary to not only observe quantum effects but also (with additional circuitry) enable the realization of a number of important goals. For continuous measurements, these include detection of zero-point motion¹, evading quantum back-action²² or Heisenberg-limited detection of displacement or force¹, preparation of non-classical photon states through motional effects³, and possibly a direct quantum non-demolition x^2 energy measurement¹⁵. Discrete measurements would enable the preparation of complex multi-phonon states²⁶, long-lived

in the transmission at a frequency $\omega_d + \Omega'_m$. **c**, When the drive amplitude is increased ($n_d \approx 10^4$), the two resonances show an avoided crossing between the eigenmodes of the coupled system. **d**, For the largest amplitude drive ($n_d \approx 5 \times 10^6$), the driven transmission spectra are split by much more than κ or Γ_m for a broad range of detunings.

quantum memory, a quantum bus architecture for quantum information processing analogous to that achieved with ion traps, and a way to investigate decoherence for large, spatially separated superposition states²⁷.

Last, we consider ground-state cooling, which is the next step forward. Assuming thermal equilibrium at a base cryostat temperature of 20 mK, we expect the cavity to be in its ground state and the thermal occupancy of the mechanical mode to be $n_m \approx 40$ quanta. Strong coupling implies that resolved sideband cooling can be used to reduce the occupancy of the mechanical mode by a factor of $\kappa/\Gamma_m \approx 5,000$ (refs 7–9), placing both (coupled) resonators in their ground state. Once this is achieved, the thermal lifetime of a single phonon would be $1/\Gamma_{th} \approx 130 \mu s$, which is orders of magnitude longer than typical superconducting qubit lifetimes. Verifying this achievement would require either a better cryogenic amplifier^{21,28} (for continuous measurements) or a superconducting qubit⁴ (for discrete measurements). Continuous measurement with microwave parametric amplifiers, which achieve nearly quantum-limited sensitivity, reduce by a factor of 1,000 the integration time necessary to resolve mechanical occupancy below the single phonon level²¹. Alternatively, if a qubit were strongly coupled to the microwave cavity^{4,26}, a single photon could be

exchanged with the qubit over a time of ~ 10 ns and would survive there for nearly $1/\kappa_0 \approx 4 \mu\text{s}$. For the largest coupling strengths achieved here, the microwave photon could then be converted into a phonon (and vice versa) in approximately $\pi/2g \approx 0.5 \mu\text{s}$, which is adequate for building up multi-phonon states^{26,29}. Beyond this, proposals exist that would allow the mechanical drum to convert microwave photon states to optical photon states, which can be used to transmit quantum information at room temperature³⁰.

Received 31 October 2010; accepted 31 January 2011.

1. Braginsky, V. B. & Khalili, F. Y. *Quantum Measurement* (Cambridge Univ. Press, 1992).
2. Mancini, S., Man'ko, V. I. & Tombesi, P. Ponderomotive control of quantum macroscopic coherence. *Phys. Rev. A* **55**, 3042–3050 (1997).
3. Bose, S., Jacobs, K. & Knight, P. L. Preparation of nonclassical states in cavities with a moving mirror. *Phys. Rev. A* **56**, 4175–4186 (1997).
4. O'Connell, A. D. et al. Quantum ground state and single-phonon control of a mechanical resonator. *Nature* **464**, 697–703 (2010).
5. Kippenberg, T. J. & Vahala, K. J. Cavity optomechanics: back-action at the mesoscale. *Science* **321**, 1172–1176 (2008).
6. Marquardt, F. & Girvin, S. M. Optomechanics. *Physics* **2**, 40 (2009).
7. Marquardt, F., Chen, J. P., Clerk, A. A. & Girvin, S. M. Quantum theory of cavity-assisted sideband cooling of mechanical motion. *Phys. Rev. Lett.* **99**, 093902 (2007).
8. Wilson-Rae, I., Nooshi, N., Zwerger, W. & Kippenberg, T. J. Theory of ground state cooling of a mechanical oscillator using dynamical backaction. *Phys. Rev. Lett.* **99**, 093901 (2007).
9. Dobrindt, J. M., Wilson-Rae, I. & Kippenberg, T. J. Parametric normal-mode splitting in cavity optomechanics. *Phys. Rev. Lett.* **101**, 263602 (2008).
10. Agarwal, G. S. & Huang, S. Electromagnetically induced transparency in mechanical effects of light. *Phys. Rev. A* **81**, 041803 (2010).
11. Weis, S. et al. Optomechanically induced transparency. *Science* **330**, 1520–1523 (2010).
12. Ekinci, K. L. & Roukes, M. L. Nanoelectromechanical systems. *Rev. Sci. Instrum.* **76**, 061101 (2005).
13. Diedrich, F., Bergquist, J. C., Itano, W. M. & Wineland, D. J. Laser cooling to the zero-point energy of motion. *Phys. Rev. Lett.* **62**, 403–406 (1989).
14. Gröblacher, S., Hammerer, K., Vanner, M. R. & Aspelmeyer, M. Observation of strong coupling between a micromechanical resonator and an optical cavity field. *Nature* **460**, 724–727 (2009).
15. Thompson, J. D. et al. Strong dispersive coupling of a high-finesse cavity to a micromechanical membrane. *Nature* **452**, 72–75 (2008).
16. Braginsky, V. B., Manukin, A. B. & Tikhonov, M. Y. Investigation of dissipative ponderomotive effects of electromagnetic radiation. *Sov. Phys. JETP* **31**, 829–830 (1970).
17. Linthorne, N. P., Veitch, P. J. & Blair, D. G. Interaction of a parametric transducer with a resonant bar gravitational radiation detector. *J. Phys. D* **23**, 1–6 (1990).
18. Regal, C. A., Teufel, J. D. & Lehnert, K. W. Measuring nanomechanical motion with a microwave cavity interferometer. *Nature Phys.* **4**, 555–560 (2008).
19. Teufel, J. D., Harlow, J. W., Regal, C. A. & Lehnert, K. W. Dynamical backaction of microwave fields on a nanomechanical oscillator. *Phys. Rev. Lett.* **101**, 197203 (2008).
20. Rocheleau, T. et al. Preparation and detection of a mechanical resonator near the ground state of motion. *Nature* **463**, 72–75 (2010).
21. Teufel, J. D., Donner, T., Castellanos-Beltrán, M. A., Harlow, J. W. & Lehnert, K. W. Nanomechanical motion measured with an imprecision below that at the standard quantum limit. *Nature Nanotechnol.* **4**, 820–823 (2009).
22. Hertzberg, J. B. et al. Back-action-evading measurements of nanomechanical motion. *Nature Phys.* **6**, 213–217 (2010).
23. Cicak, K. et al. Low-loss superconducting resonant circuits using vacuum-gap-based microwave components. *Appl. Phys. Lett.* **96**, 093502 (2010).
24. Blencowe, M. P. & Buks, E. Quantum analysis of a linear dc squid mechanical displacement detector. *Phys. Rev. B* **76**, 014511 (2007).
25. Boller, K.-J., Imamolu, A. & Harris, S. E. Observation of electromagnetically induced transparency. *Phys. Rev. Lett.* **66**, 2593–2596 (1991).
26. Hofheinz, M. et al. Synthesizing arbitrary quantum states in a superconducting resonator. *Nature* **459**, 546–549 (2009).
27. Marshall, W., Simon, C., Penrose, R. & Bouwmeester, D. Towards quantum superpositions of a mirror. *Phys. Rev. Lett.* **91**, 130401 (2003).
28. Castellanos-Beltrán, M. A., Irwin, K. D., Hilton, G. C., Vale, L. R. & Lehnert, K. W. Amplification and squeezing of quantum noise with a tunable Josephson metamaterial. *Nature Phys.* **4**, 929–931 (2008).
29. Akram, U., Kiesel, N., Aspelmeyer, M. & Milburn, G. J. Single-photon optomechanics in the strong coupling regime. *N. J. Phys.* **12**, 083030 (2010).
30. Regal, C. A. & Lehnert, K. W. From cavity electrodynamics to cavity optomechanics. *J. Phys. Conf. Ser.* **264**, 012025 (2011).

Acknowledgements We thank A. W. Sanders for taking the micrograph in Fig. 1b, and acknowledge discussions with T. Donner, J. H. Harlow and K. W. Lehnert. This paper is a contribution by the National Institute of Standards and Technology and not subject to US copyright.

Author Contributions J.D.T. and R.W.S. conceived the device. J.D.T. designed the circuit. J.D.T. and D.L. fabricated the devices. J.D.T. performed and analysed the measurements. R.W.S. oversaw all aspects of this work. All authors provided experimental support and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.D.T. (john.teufel@nist.gov).

Permanent El Niño during the Pliocene warm period not supported by coral evidence

Tsuyoshi Watanabe¹, Atsushi Suzuki², Shoshiro Minobe¹, Tatsunori Kawashima¹, Koji Kameo³, Kayo Minoshima², Yolanda M. Aguilar⁴, Ryoji Wani⁵, Hodaka Kawahata⁶, Kohki Sowa¹, Takaya Nagai¹ & Tomoki Kase⁷

The El Niño/Southern Oscillation (ENSO) system during the Pliocene warm period (PWP; 3–5 million years ago) may have existed in a permanent El Niño state with a sharply reduced zonal sea surface temperature (SST) gradient in the equatorial Pacific Ocean¹. This suggests that during the PWP, when global mean temperatures and atmospheric carbon dioxide concentrations were similar to those projected for near-term climate change², ENSO variability—and related global climate teleconnections—could have been radically different from that today. Yet, owing to a lack of observational evidence on seasonal and interannual SST variability from crucial low-latitude sites, this fundamental climate characteristic of the PWP remains controversial^{1,3–10}. Here we show that permanent El Niño conditions did not exist during the PWP. Our spectral analysis of the $\delta^{18}\text{O}$ SST and salinity proxy, extracted from two 35-year, monthly resolved PWP *Porites* corals in the Philippines, reveals variability that is similar to present ENSO variation. Although our fossil corals cannot be directly compared with modern ENSO records, two lines of evidence suggest that Philippine corals are appropriate ENSO proxies. First, $\delta^{18}\text{O}$ anomalies from a nearby live *Porites* coral are correlated with modern records of ENSO variability. Second, negative- $\delta^{18}\text{O}$ events in the fossil corals closely resemble the decreases in $\delta^{18}\text{O}$ seen in the live coral during El Niño events. Prior research advocating a permanent El Niño state may have been limited by the coarse resolution of many SST proxies, whereas our coral-based analysis identifies climate variability at the temporal scale required to resolve ENSO structure firmly.

There is still considerable debate over the characteristics of ENSO during the PWP, even though several previous models and geochemical proxy-based reconstructions have been generated^{1,3–10}. The clearest controversy has appeared in recent publications in which two teams using the same geochemical approaches with the same sediment cores from western and eastern Pacific sites obtained opposing results for the climatic conditions related to ENSO during the PWP in the Pacific Ocean. Both studies combined Mg/Ca and $\delta^{18}\text{O}$ ratios for the same species of planktonic foraminifera to reconstruct the western and eastern gradients of SST and salinity. These factors constrain the background climate state and the variability of ENSO occurrence. The first-presented work⁵ on a strong SST gradient between the western and eastern equatorial Pacific concluded that a La Niña-like state was dominant throughout the PWP. Shortly thereafter, it was suggested¹, on the basis of Mg/Ca and $\delta^{18}\text{O}$ data with higher temporal resolution from the same set of sediment cores, that there were permanent El Niño-like conditions during the PWP. In a warm and wet climatic background with global warming and freshening, the larger and fresher water mass that permanently exists in the eastern Pacific Ocean might prevent periodic occurrences of El Niño, or the newer and stronger heat conditions might lead to similar, stronger and more frequent El Niño

events. Reconstructing ENSO characteristics in the PWP is important for understanding future climate change scenarios. Both permanent El Niño- and La Niña-like conditions have been found in alkenone-based SST reconstructions^{9,10} and model-based estimations^{6,7}. Owing to the limited temporal resolution of ocean sediment cores, however, no direct evidence exists about the variation of ENSO phenomena, which have seasonal and interannual variabilities caused by interactions between the ocean and atmosphere.

Today ENSO is defined as a shifting of the western Pacific warm pool (WPWP), the warmest surface sea water in the western Pacific, towards the eastern Pacific in 3–7-yr intervals¹¹. This warm water mass becomes thicker in the eastern Pacific and prevents upwellings in this region during El Niño conditions. Simultaneously, in the atmosphere the Hadley cell and Walker circulation shift towards the eastern Pacific side, resulting in cooler and drier conditions on the western side of the Pacific Ocean.

Coral skeletons in tropical and subtropical regions hold a wealth of information about past climate changes with monthly temporal resolution from the geochemical signatures of their annual bands^{12–14}. The $\delta^{18}\text{O}$ ratio in coral skeletons has been widely used to reconstruct past SSTs and water balance^{12–15}. The utility of skeletal $\delta^{18}\text{O}$ from the coral genus *Porites* in monitoring interannual climate fluctuations in the tropical Pacific region has been widely demonstrated for both modern¹¹ and fossil corals^{13,14}. The effect of ENSO on *Porites* $\delta^{18}\text{O}$ records in the WPWP region has also been evident in previous studies on living and fossil specimens^{14,15}. During El Niño events, the combined effect of lower SST and reduced monsoon rainfall (higher $\delta^{18}\text{O}$ values of sea water ($\delta^{18}\text{O}_{\text{w}}$)) has served to increase $\delta^{18}\text{O}$ in *Porites*, in association with prevailing higher monsoon rainfall (lower $\delta^{18}\text{O}_{\text{w}}$ values). El Niño events are triggered when eastern equatorial Pacific SSTs are unusually high, and these events are typically associated with cooler SSTs¹⁶ and below-average monsoon rainfall¹⁷ in the southwestern North Pacific. In contrast, changes in SSTs in the southwestern North Pacific are less marked during La Niña events (when the eastern equatorial Pacific is cooler than average), but rainfall is enhanced owing to more vigorous summer monsoons. At the PWP fossil coral sampling site in the Philippines, the interannual climate characteristics of SSTs (Supplementary Fig. 1a) and precipitation¹⁸ (Supplementary Fig. 1b) observed during El Niño events were clearly evident. Although the response time of corals to ENSO was slightly different for SST and precipitation, a modern coral $\delta^{18}\text{O}$ record near the fossil coral site captured recent El Niño signals (Fig. 1; see Supplementary Information for a more detailed discussion). These findings suggest that our coral site and coral records were adequate for detecting ENSO events.

To decipher the seasonal and interannual characteristics of the Pliocene ENSO, we used two 35-yr $\delta^{18}\text{O}$ records with monthly resolution deduced from fossil *Porites* corals. We found these exceptionally well-preserved fossil corals in slightly separated stratigraphic levels of

¹Department of Natural History Sciences, Faculty of Science, Hokkaido University, Sapporo 060-0810 Japan. ²Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology, Tsukuba 305-8567, Japan. ³Department of Earth Sciences, Faculty of Science, Chiba University, Chiba 263-8522, Japan. ⁴Petrolab, Mines and Geosciences Bureau, North Avenue, Diliman, Quezon City 1101, Philippines. ⁵Interdisciplinary Research Center, Yokohama National University, Yokohama 240-8501, Japan. ⁶Atmosphere and Ocean Research Institute, The University of Tokyo, Kashiwa 277-5564, Japan. ⁷Department of Geology and Paleontology, National Museum of Nature and Science, Tokyo 169-0073, Japan.

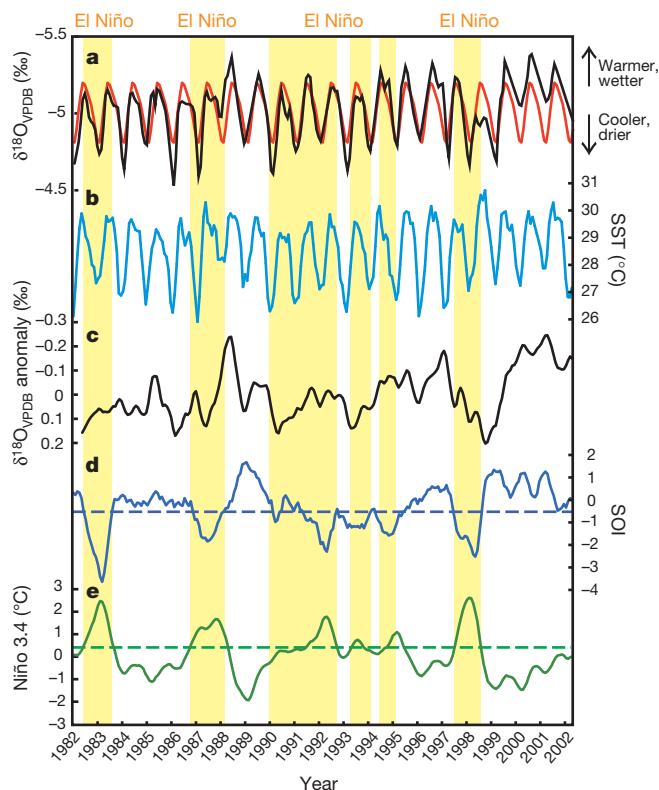


Figure 1 | Comparison of modern coral $\delta^{18}\text{O}$, local SST and coral $\delta^{18}\text{O}$ anomalies and ENSO proxy time series. Monthly resolution $\delta^{18}\text{O}$ time series (a; black) and climatology (a; red) are shown in comparison with monthly SSTs (b), modern coral $\delta^{18}\text{O}$ anomalies (c), the Southern Oscillation index (SOI; d) and the temperature anomaly in the Niño 3.4 region (e). ($\delta^{18}\text{O} = (^{18}\text{O}/^{16}\text{O})_{\text{sample}} / (^{18}\text{O}/^{16}\text{O})_{\text{standard}} - 1$; here the standard is Vienna Pee Dee Belemnite (VPDB)). Monthly SST data came from the National Oceanic and Atmospheric Administration²³ (NOAA OI SST version 2). The yellow shading indicates the time periods of recent El Niño events, defined as those periods when the temperature anomaly in the Niño 3.4 region is above 0.5°C (green dotted line) and/or the SOI is less than 0.5 (blue dotted line).

the Tartaro Formation on the island of Luzon in the Philippines (Supplementary Fig. 1a, b). Observation of nannofossil assemblages revealed the ages of the fossils to be approximately 3.5–3.8 million years (age determination was based on the occurrence of *Sphenolithus* spp. and the absence of *Reticulofenestra pseudumbilicus*¹⁹ ($>7\ \mu\text{m}$)). For geochemical analysis, we selected well-preserved specimens from large coral colonies of *Porites* sp. from the coral boulders recovered.

Selected samples were examined to test for diagenetic alteration (Methods Summary and Supplementary Figs 2–4). Coral specimens containing no detectable altered mineral phases (corals 1 and 2) were used for detailed $\delta^{18}\text{O}$ measurements (Methods Summary and Supplementary Fig. 2). Selected colonies were cut into slabs (5 mm thick) parallel to the axis of maximum growth and X-rayed. The X-rays revealed distinct density band couplets (Supplementary Fig. 2). There were 35 distinct seasonal bands in each of the two time series (Fig. 2), and distinct attenuations in seasonal amplitudes were found in the $\delta^{18}\text{O}$ values for both PWP coral records. These attenuations also appeared in modern coral records during El Niño events in WPWP regions (Indonesia¹⁵, Papua New Guinea¹⁴, New Caledonia¹⁵ and the Philippines; Fig. 1). The seasonal amplitudes of skeletal $\delta^{18}\text{O}$ values for coral 1 were significantly larger than those for coral 2, suggesting that these two coral colonies grew for different time periods during the PWP. These findings suggest that ENSO existed during different time windows within the PWP.

We calculated the power spectral densities for the two time series of anomalies from seasonal cycles for corals 1 and 2 (Fig. 3a), as well as those for a modern coral record (Fig. 3b) and for the temperature anomaly in the Niño 3.4 region over the same intervals (1950–1984 and 1985–2010, respectively; Fig. 3c). We also determined the average power spectrum for both time series (Supplementary Fig. 5 and Supplementary Information). Local peaks with a 3.5-yr period were commonly found in all spectra. Although the peak–trough contrast was relatively small for coral 1 and the individual spectra of corals 1 and 2 were different in character, such fractionation was also found in recent ENSO spectra (Fig. 3c) and in the reconstructed ENSO variability during the last millennium²⁰ and on the glacial–interglacial timescale¹⁴. The power spectral density of the PWP coral records suggests that ENSO-like interannual variability occurred over the tropical Pacific in the Pliocene period.

The early PWP (3–4.5 million years ago) has been widely studied as an analogue of future global climate changes that human beings will face in the coming decades. It is therefore crucial to determine the state of ENSO during the PWP and to predict how it will behave in the coming centuries in terms of strength and frequency. Until now, the characteristics of the interannual variability of ENSO during the PWP have been unknown. A recent modelling study⁷ using a coupled ocean–atmosphere general circulation model has suggested that El Niño events occurred in the Pliocene. Our results provide the first proxy evidence of ENSO-like interannual variability occurring during the PWP. This direct evidence could be used to determine the mechanism underlying ENSO, to predict future climate behaviour, for example to assess how a warm climate affects the variability and amplitude of ENSO. Century-long coral records have been used to discuss the possibility that the strength of ENSO variability may depend on the

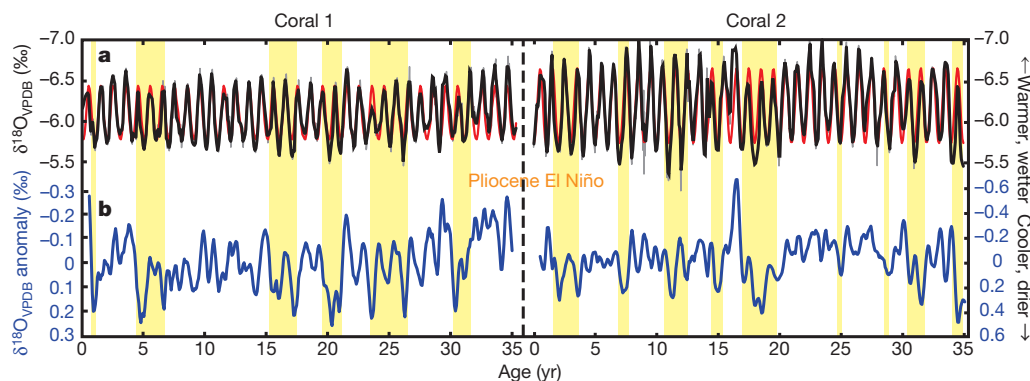


Figure 2 | Pliocene El Niño recorded in PWP oxygen isotopic records. a, Oxygen isotopic time series (black), best-fit climatology (red) and raw data plots (grey) for PWP corals (corals 1 and 2). Oxygen isotope records were interpolated to a resolution of 12 samples per year and climatology was

calculated using annual and semi-annual sinusoidal functions. b, Anomalies in the PWP coral $\delta^{18}\text{O}$ time series ($\delta^{18}\text{O}$ minus climatology in a). The standard deviations of the anomalies were 0.139 for coral 1 and 0.217 for coral 2. The yellow shading indicates possible Pliocene El Niño events.

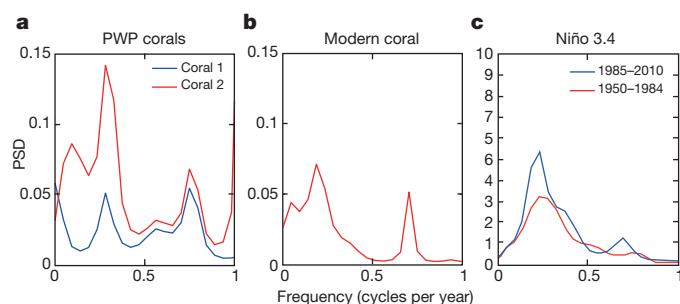


Figure 3 | Power spectral densities. Power spectral densities (PSDs) estimated by the maximum-entropy method for PWP fossil corals (**a**; blue, coral 1; red, coral 2), modern coral $\delta^{18}\text{O}$ (**b**) and the Niño 3.4 index (**c**; red, 1950–1984; blue, 1975–2010). A local peak at around 0.3 cycles per year is common to all of the spectra. This peak was statistically significantly larger than the trough at around 0.5 cycles per year at the 90% confidence level for coral 2. The peak–trough contrast was too small to be significant, however, for coral 1.

average conditions of factors such as SST²¹. Our PWP coral records suggest that the characteristics of the Pliocene ENSO were similar to those of recent ENSO events, and stand in contrast to studies suggesting that past and future warm climate background conditions will lead to a permanent El Niño state^{1,6,8,9}. The present study reveals active ENSO dynamics resulting from the interannual variability of WPWP movements during the PWP, supporting the hypothesis that stronger winds in the tropical Pacific²² act as a possible driving force behind ENSO when SSTs are higher than present levels⁴. This stronger wind enhanced by a heated ocean during the PWP probably led to a positive-feedback effect on ENSO dynamics, and such positive-feedback climate systems may occur in the future. That said, on the basis of millennium-long records of overlapping young and fossil corals it has been implied that the internal dynamics in the ENSO system itself might also cause the large and abrupt changes in ENSO characteristics even without external forcing²⁰. ENSO might have a stronger or more frequent impact on climates in the mid to high latitudes in the future. Additional work is necessary to understand the role of the Pliocene El Niño on the teleconnection pattern in the Pacific Ocean during warm conditions.

METHODS SUMMARY

Screening methods for coral diagenesis. We screened and selected two exceptionally well-preserved fossil specimens for signs of diagenetic alteration using the following methods: (1) screening tests, including X-ray radiography and X-ray diffraction analysis, and (2) microanalysis of thin sections using the high-energy synchrotron X-ray diffraction analysis in combination with microstructural observation by scanning electronic microscopy and optical microscopic observation (see Supplementary Information for more detailed methods and discussion of diagenesis).

Oxygen isotope analysis. We performed microsampling along the major growth axis at 400- μm intervals. Microsamples of two coral cores (corals 1 and 2), which each weighed approximately 70–110 μg , were reacted with 100% H_3PO_4 at 90 °C in an automated carbonate preparation device coupled to a Micromass Optima mass spectrometer. Isotopic data are reported as per mil (‰) deviations relative to Vienna PeeDee Belemnite. The typical precision was better than 0.1‰ (see Supplementary Information for more details).

Received 28 June; accepted 22 December 2010.

1. Wara, M. W., Ravelo, A. C. & Delaney, M. L. Permanent El Niño-like conditions during the Pliocene warm period. *Science* **309**, 758–761 (2005).

2. Mix, A. C. *et al.* Benthic foraminiferal stable isotope record from Site 849, 0–5 Ma: local and global climate changes. *Proc. ODP Sci. Res.* **138**, 371–412 (1995).
3. Molnar, P. & Cane, M. A. El Niño's tropical climate and teleconnections as a blueprint for pre-Ice Age climates. *Paleoceanography* **17**, 1021 (2002).
4. Ravelo, A. C., Andreasen, D. H., Lyle, M., Lyle, A. O. & Wara, M. W. Regional climate shifts caused by gradual global cooling in the Pliocene epoch. *Nature* **429**, 263–267 (2004).
5. Rickaby, R. E. M. & Halloran, A. P. Cool La Niña during the warmth of the Pliocene? *Science* **307**, 1948–1952 (2005).
6. Fedorov, A. V. *et al.* The Pliocene paradox (mechanisms for a permanent El Niño). *Science* **312**, 1485–1489 (2006).
7. Haywood, A. M., Valdes, P. J. & Peck, V. L. A permanent El Niño-like state during the Pliocene? *Paleoceanography* **22**, PA1213 (2007).
8. Molnar, P. & Cane, M. A. Early Pliocene (pre-Ice Age) El Niño-like global climate: which El Niño? *Geosphere* **3**, 337–365 (2007).
9. Lawrence, K. T., Liu, Z. & Herbert, T. D. Evolution of the eastern tropical Pacific through Plio-Pleistocene glaciation. *Science* **312**, 79–83 (2006).
10. Dowsett, H. J. & Robinson, M. M. Mid-Pliocene equatorial Pacific sea surface temperature reconstruction: a multi-proxy perspective. *Phil. Trans. R. Soc. A* **367**, 109–125 (2009).
11. Philander, S. G. H. *El Niño, La Niña and the Southern Oscillation* 9–57 (Academic, 1990).
12. Cole, J. E., Fairbanks, R. G. & Shen, G. T. Recent variability in the southern oscillation: isotopic results from Tarawa Atoll coral. *Science* **260**, 1790–1793 (1993).
13. Corrège, T. *et al.* Interdecadal variation in the extent of South Pacific tropical waters during the Younger Dryas event. *Nature* **428**, 927–929 (2004).
14. Tudhope, A. W. *et al.* Variation in the El Niño–Southern Oscillation through a glacial–interglacial cycle. *Science* **291**, 1511–1517 (2001).
15. Watanabe, T. *et al.* Oxygen isotope systematics in *Diploastrea heliopora*: new coral archive of tropical paleoclimate. *Geochim. Cosmochim. Acta* **67**, 1349–1358 (2003).
16. Rasmusson, E. M., & Carpenter, T. H. Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Mon. Weath. Rev.* **110**, 354–384 (1982).
17. Ropelewski, C. F. & Halpert, M. S. Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Weath. Rev.* **115**, 1606–1626 (1987).
18. Dai, A. & Wigley, T. M. L. Global patterns of ENSO-induced precipitation. *Geophys. Res. Lett.* **27**, 1283–1286 (2000).
19. Raffi, I. *et al.* A review of calcareous nannofossil astrobiochronology encompassing the past 25 million years. *Quat. Sci. Rev.* **25**, 3113–3137 (2006).
20. Cobb, K. M., Charles, C. D., Cheng, H. & Edwards, R. L. El Niño/Southern oscillation and tropical Pacific climate during the last millennium. *Nature* **424**, 271–276 (2003).
21. Urban, F. E., Cole, J. E. & Overpeck, J. T. Influence of mean climate change on climate variability from a 155-year tropical Pacific coral record. *Nature* **407**, 989–993 (2000).
22. Fedorov, A. V., Brierley, C. M. & Emanuel, K. Tropical cyclones and permanent El Niño in the early Pliocene epoch. *Nature* **463**, 1066–1070 (2010).
23. Reynolds, R. *et al.* Daily high-resolution blended analyses for sea surface temperature. *J. Clim.* **20**, 5473–5496 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank H. H. Ramos for a research permit and W. Mago and E. Azurin for their help with fieldwork. Special thanks go to F. P. Siringan for the study of modern coral; H. Nomura and K. Nakamura for preparing thin sections; R. Miyawaki for assistance with X-ray diffraction; S. Motai, Y. Seto and K. Omori for their assistance with synchrotron X-ray diffraction; C. Shimada, M. Ikeda and K. Hagino-Tomioka for their scanning electron microscopy observations; Y. Yoshinaga for microsampling; and A. Yamazaki for preparing the figures.

Author Contributions T.W., A.S. and T. Kase designed the research. T. Kawashima and K.M. performed isotopic analysis under the supervision of T.W., A.S. and H.K. S.M. performed statistical analysis and climatology. K.K. observed nannofossil assemblages. T.W., T. Kawashima, Y.M.A., R.W. and K.S. carried out the field survey under the supervision of T. Kase. T.W. and K.S. performed synchrotron X-ray diffraction analysis under the supervision of T.N. T.W. wrote the paper in discussion with all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.W. (nabe@mail.sci.hokudai.ac.jp).

A rapid mechanism to remobilize and homogenize highly crystalline magma bodies

Alain Burgisser¹ & George W. Bergantz²

The largest products of magmatic activity on Earth, the great bodies of granite and their corresponding large eruptions, have a dual nature: homogeneity at the large scale and spatial and temporal heterogeneity at the small scale^{1–4}. This duality calls for a mechanism that selectively removes the large-scale heterogeneities associated with the incremental assembly⁴ of these magmatic systems and yet occurs rapidly despite crystal-rich, viscous conditions seemingly resistant to mixing^{2,5}. Here we show that a simple dynamic template can unify a wide range of apparently contradictory observations from both large plutonic bodies and volcanic systems by a mechanism of rapid remobilization (unzipping) of highly viscous crystal-rich mushes. We demonstrate that this remobilization can lead to rapid overturn and produce the observed juxtaposition of magmatic materials with very disparate ages and complex chemical zoning. What distinguishes our model is the recognition that the process has two stages. Initially, a stiff mushy magma is reheated from below, producing a reduction in crystallinity that leads to the growth of a subadjacent buoyant mobile layer. When the thickening mobile layer becomes sufficiently buoyant, it penetrates the overlying viscous mushy magma. This second stage rapidly exports homogenized material from the lower mobile layer to the top of the system, and leads to partial overturn within the viscous mush itself as an additional mechanism of mixing. Model outputs illustrate that unzipping can rapidly produce large amounts of mobile magma available for eruption. The agreement between calculated and observed unzipping rates for historical eruptions at Pinatubo and at Montserrat demonstrates the general applicability of the model. This mechanism furthers our understanding of both the formation of periodically homogenized plutons (crust building) and of ignimbrites by large eruptions.

There is general agreement that magmatic systems in all tectonic settings are open to the input of heat and mass, and are incrementally assembled⁴. This is manifested in complex crystal zoning patterns, age diagnostics and volatile degassing budgets^{2,3,6,7}, which all show that the magmatic systems experienced periods of rejuvenation by chemical and thermal fluctuations. Yet despite diverse inputs, many of the largest magma bodies (frozen as large granite plutons) and their erupted products have a uniform bulk composition at meso-to-macro scales¹ even though adjacent crystals have different histories^{2,6}. They also rarely preserve a record of the assembly process, despite the evidence that they have formed incrementally and for the entrainment of older magmatic materials⁴. Hence, some process must be capable of efficiently removing the large-scale heterogeneities associated with assembly, and creating an environment of common intensive variables despite the viscous nature of these crystal-rich systems. To complicate any simple mechanical model that accounts for both incremental assembly and homogenization, the geological and geophysical evidence requires that large magmatic systems persist as long-lived mushes, which are seemingly resistant to homogenization^{6,8,9}. And although gas sparging¹⁰ and self-mixing¹¹ have been proposed as mechanisms of rejuvenation or mixing, neither is fully consistent with the observations that

magmatic systems may not be subject to regular, substantial volatile throughput and that they spend long periods as rheologically stiff mushes inhibiting simple convection¹².

On the basis of petrologic evidence, Mahood¹³ offered one solution by proposing a process of ‘defrosting’, in which material that is mechanically locked into the solidified margins of a magma body could be liberated by melting, and then back-mixed into the more fluid core. Although theoretical modelling¹⁴ of such melting has been carried out, recent advances in the study of the rheology of crystal-rich magmas drive us to reconsider some aspects of this modelling (Supplementary Methods). Briefly, the two main controls of bulk viscosity, melt water content and crystal content, have opposite effects that closely compensate each other¹⁵. For example, the viscosity of magma with rhyolitic melt remains remarkably stable at around 10^4 Pa s over most of its *in situ* crystallization, until a critical crystal concentration is reached above which the viscosity significantly increases to produce the mush state. We reassess the notion of a ‘defrosting’ front by casting it as a remobilization front moving into a mushy core, and show that it is a process that can act rapidly to rejuvenate magma mushes, and to mix magmatic materials of diverse ages and character into a near-uniform state.

Our model considers the fate of a magma reservoir filled with a highly crystalline mush that is subjected to reheating from below by a fresh magma intrusion. The melting of the mush by the new intrusion causes the dismantlement of the crystal framework, which frees the mush little by little to form a mobile, more melt-rich and less dense layer (Fig. 1). If the melting continues undisturbed, the pre-existing mush becomes entirely remobilized as the mobile layer fills the entire chamber. We call this process, which corresponds to the classic approach¹⁴, ‘stable front remobilization’. A novel consequence of the stair-step behaviour of the mush viscosity is that the hot, mobile layer is buoyant with respect to the colder mush. The thickening mobile layer is thus more and more prone to Rayleigh–Taylor instabilities that might penetrate the overlying mush. We call this process of penetrative overturn ‘unzipping’ (meaning rapid remobilization by an unstable front).

Figure 1 shows the evolution of the thickness of the mobile layer as a function of time when all model parameters are set to typical values for a mid-crustal reservoir (Table 1). Shortly after the emplacement of the basal, hot intrusion, the growth of the mobile layer is purely conductive. After 1.8 days, the mobile layer starts convecting to finally reach the full chamber thickness after about 95 years, very close to the 101 years given by a previous model¹⁴. However, if the possibility of Rayleigh–Taylor instabilities at the interface is considered, these instabilities will grow faster than the mobile layer after only 68 days, when the layer is 9 m thick. At this point the buoyant mobile layer forms an ascending plume that penetrates the mush and reaches the chamber roof 75 days later, causing partial overturn of the remaining mush and leaving the chamber in a remobilized state. During that penetration time, the interface between mush and mobile layer would have steadily moved about 8 m.

Fixing all parameters to their default value (Table 1) except for the mush viscosity yields results shown as dashed lines on Fig. 2. Under

¹Institut des Sciences de la Terre d’Orléans, CNRS/INSU, Université d’Orléans, Université François Rabelais—Tours, 1A rue de la Férolière, 45071 Orléans cedex 2, France. ²Department of Earth and Space Sciences, University of Washington, Box 351310, Seattle, Washington 98195-1310, USA.

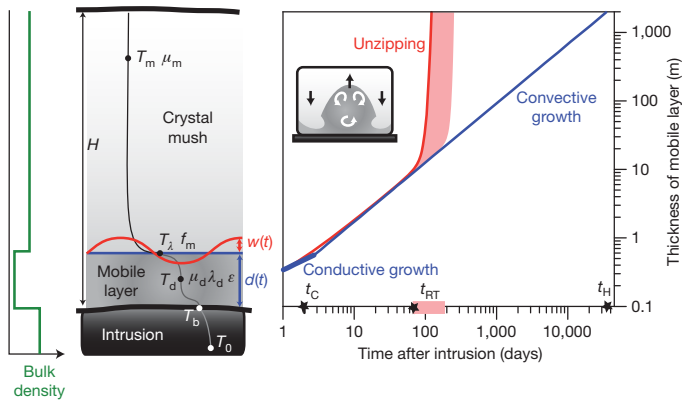


Figure 1 | Schematic of a stagnant mid-crustal reservoir being reheated from below by an intrusion. The mush reheats in two stages by forming a convecting mobile layer that grows steadily before becoming unstable and eventually overturning the remaining mush (inset), a process we call unzipping. The process is driven by buoyancy, as illustrated by the schematic bulk density profile on the left. Our model describes the temporal evolution of the mobile layer thickness d and the interface instability amplitude w , as a function of ten free parameters (see Table 1 for definitions). On the right is a model output for a typical mid-crustal reservoir. The unstable front ($w + d$) is shown in red and the stable front (d) is shown in blue. The horizontal axis indicates the time at which the layer starts convecting t_c , the time at which Rayleigh–Taylor instabilities grow faster than the mobile layer t_{RT} , and the time at which the mobile layer would fill the chamber in the absence of Rayleigh–Taylor instabilities t_H . The time interval caused by the uncertainty of the mush rheology is indicated by pink shading, and the minimum intrusion thickness for unzipping to occur is 1.9 ± 0.8 m (Supplementary Discussion).

such conditions, unzipping takes between a couple of months and a couple of decades to start and this period is always shorter than for stable front remobilization. We varied the other nine parameters within the range of values expected for mid-crustal magma bodies (Table 1). The onset of convection within the mobile layer is mostly controlled by the intrusion temperature and the layer viscosity, and occurs between a few days and several months after the new intrusion (Fig. 2). Stable front remobilization is most affected by the intrusion temperature, mush thickness and mush viscosity. In agreement with previous findings¹⁴, it occurs between a century and more than 10,000 years after the new intrusion. The strongest controls of the unzipping onset time are mush viscosity, mobile layer viscosity, and

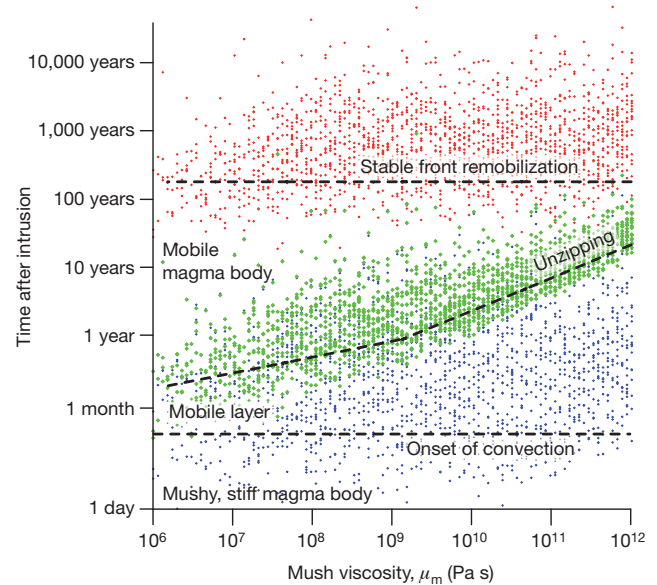


Figure 2 | Time for mush melting from a steady source of heat as a function of mush viscosity. The dashed lines correspond to a typical system (Table 1), with the first line marking the onset of convection within the mobile layer, the second marking the onset time of unzipping, and the last line indicating the time taken by stable front remobilization to reach the reservoir roof. Rejuvenation happens either by unzipping or stable front remobilization, whichever occurs first. Coloured dots (blue, t_c ; green, t_{RT} ; red, t_H) illustrate the variations (10^5 Monte Carlo runs, only 1 out of 5 shown) of these three timescales as a function of the nine free parameters listed in Table 1. The kink on the unzipping curve at a mush viscosity of $10^{9.1}$ Pa s is caused by limiting the instability wavelength at the reservoir height so as to avoid unrealistically large instabilities (Supplementary Methods).

intrusion temperature, owing to their wide natural variation. The other parameters have only minor effects on unzipping, causing the onset time to vary within a factor of less than 1.9 of the standard value (Supplementary Discussion). We conclude that all but the stiffest mid-crustal reservoirs are subjected to fast remobilization by unzipping within a time frame of as little as a few months, but always less than a few centuries. These durations are much shorter than those related to other mechanisms^{14,16,17} and are only comparable to the conductive remobilization of accumulated intrusions quenched to glass¹⁸.

The thicknesses of intrusions needed to trigger unzipping lie between 0.2 and 83 m (Supplementary Discussion), which is consistent with sills observed at the roots of plutons¹⁹. Using unzipping times, these thicknesses correspond to magma supply rates of between 2.1×10^{-3} and $0.125 \text{ km}^3 \text{ yr}^{-1}$, which is in agreement with the $(1.2\text{--}6) \times 10^{-2} \text{ km}^3 \text{ yr}^{-1}$ estimated for the flare-up of a major ignimbrite province²⁰. This is in stark contrast with models ignoring the buoyancy of the mobile layer^{12,14}, which require intrusion thicknesses of the same order as that of the mush. Only a very small amount of basalt cooling, by 100°C or so, is sufficient to create the modest melting of the thin mobile layer at the base of the mush, because our model involves only partial mush melting (typically 20 vol.%, Table 1). In the typical case (Fig. 1), a 3-m basalt sill would be reaching 60% crystallinity in 70 days (ref. 14), but unzipping takes only 68 days to occur. Any larger intrusion has a thermal history decoupled from that of the mush. The small amount of intrusive material needed and the rapidity of the process leads us to consider that unzipping is an easily triggered mechanism yielding large volumes of eruptible magmas. Unzipping is likely to repeat itself throughout the life of the mush, until its eruption or thermal death, because cooling magma bodies spend most of their existence as mushes⁹. Only a few successive overturns are enough to homogenize the magma body⁹. This leads us to view unzipping as a phenomenon that probably allows for the incremental growth of plutons and their periodic homogenization^{4,6}.

Table 1 | Model parameters for mid-crustal reservoirs

	Minimum	Maximum	Typical value
Free parameters			
ε , melt volume fraction in mobile layer	0.5	0.8	0.6
f_m , volume fraction of melting during mush reheating	0.1	0.4	0.2
T_0 , temperature of the intrusion ($^\circ\text{C}$)	870	1,200	1,100
T_m , initial temperature of mush ($^\circ\text{C}$)	700	825	750
T_d , temperature at mush-to-mobile transition ($^\circ\text{C}$)	730	850	800
λ_d , viscosity ratio within the mobile layer	1	10	5
μ_d , mobile layer average viscosity (Pa s)	10^3	10^7	10^4
μ_m , Newtonian mush viscosity (Pa s)	10^6	10^{12}	10^9
H , total height of the reservoir (m)	500	5,000	2,000
w_0 , initial perturbation of mush interface (m)	0.01	1	0.1
Calculated parameters			
T_b , temperature at base, average of T_0 and T_m ($^\circ\text{C}$)	785	1,013	925
T_d , mid-temperature of mobile layer ($^\circ\text{C}$)	758	947	874
λ_m , viscosity ratio between mobile layer and mush	10^{-8}	10^{-6}	10^{-5}
Mush crystallinity ($1 - \varepsilon + f_m$) (%)	50	80	60
Model output			
$d(t)$, thickness of the mobile layer (m)			
$w(t)$, amplitude of instabilities (m)			
t_c , onset time of convection in mobile layer (s)			
t_H , time for the mobile layer to reach chamber roof (s)			
t_{RT} , onset time of unzipping (s)			

See Methods for detailed model and Fig. 1 for illustration.

We tested our model against three eruptions involving the remobilization of stagnant reservoirs. The 1991 Pinatubo eruption of about 5 km³ dense rock equivalent (DRE)²¹, the current eruption of Soufrière Hills, Montserrat, exceeding 0.44 km³ DRE²², and the ~28-Myr-old Fish Canyon ignimbrite of 5,000 km³ DRE²³ involve reservoirs of widely different scales. We considered the mush viscosity as a free parameter, but constrained the other ones using petrologic studies (Table 2). This choice takes into account the uncertainty induced by the current lack of comprehensive framework to quantify mush rheology, without reducing the generality of our model (Supplementary Discussion). As a result, we cast our model predictions as the ratio between mush and mobile layer viscosities to the unzipping onset time (Fig. 3). The observed timescale between the arrival of fresh magma under the stagnant reservoir and the eruption gives an estimation of the unzipping timescale. The observed viscosity ratio is independently estimated using the full range of cold and reheated mush crystallinities and three different rheology models (Supplementary Discussion). There are no field data for the Fish Canyon ignimbrite, because the timing of the eruption and the crystallinity of the mush are unknown.

In the case of Pinatubo, the agreement between prediction and observations is reasonable, the predictions of unzipping time spanning a potential two orders of magnitude. In the case of Montserrat, the field-derived time estimates mostly overlap the ones predicted by our model. If some confidence is given to our model, the fact that the overlap is restricted to the longest timescales would imply that the mush beneath Soufrière Hills was remobilized significantly before eruption, although not as early as in the seismic crisis of 1966–67 (ref. 22). Stable front remobilization at both volcanoes would require as much time as there actually was between eruptions, which would leave an unreasonably short time for chamber replenishment and mush formation. The Fish Canyon, one of the largest ignimbrites on Earth²³, could have been remobilized in less than a couple of centuries. This duration, much shorter than the 100,000–200,000 years needed for gas sparging¹⁶, is only slightly longer than that of the much smaller magmatic system at Montserrat, which indicates that magma body size is not the main control of unzipping.

Although we assume that eruption occurs after mush remobilization, the case of Montserrat suggests that remobilization does not necessarily trigger eruption. Even if it is an efficient mechanism to generate large quantities of mobile, crystal-rich magma available for eruption, unzipping is a necessary but not sufficient condition for eruption. Our findings imply that the pre-eruptive partition between mobile

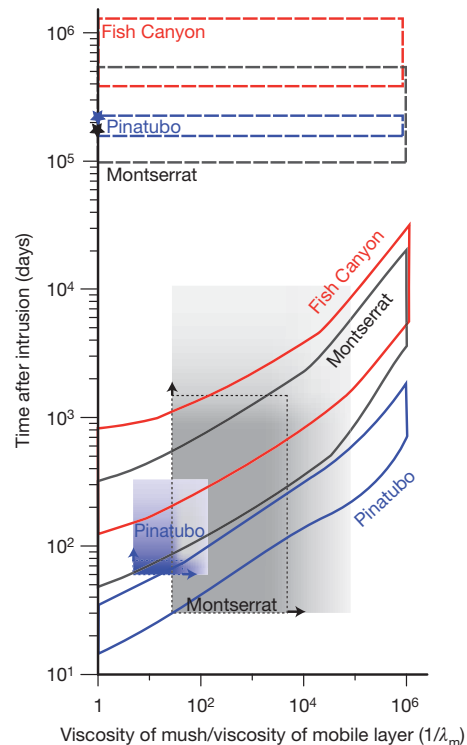


Figure 3 | Observed and predicted timescales for mush remobilization by unzipping for three natural cases. The solid-line polygons are unzipping predictions using the input parameters listed in Table 2. The coloured areas cover the ranges of the observed timescales between the arrival of fresh magma under the stagnant reservoir and the eruption (vertical axis), and the ratio between the viscosities of cold and reheated mush (horizontal axis, Table 2). Arrows and dotted rectangles delimit likely estimates and fading colour gradients indicate more speculative maximum estimates. Black and blue stars are previous eruptions at Montserrat and Pinatubo, respectively. The dashed-line polygons at the top are stable front remobilization times, which closely match earlier work¹⁴.

and stiff magmas can rapidly change during periods of volcanic unrest regardless of reservoir size and that an initially largely stagnant reservoir does not guarantee a small-scale eruption. As a result, mostly solidified magma bodies are at a crossroads between ignimbrite formation by large eruptions and crust-building by the formation of periodically homogenized plutons.

Table 2 | Natural cases of remobilized magma reservoirs

	Montserrat	Pinatubo	Fish Canyon
Model input			
ε	0.5–0.55 (0.53)*	0.85–0.74 (0.8)*	0.55–0.6 (0.55)*
f_m	0.15	0.27	0.2–0.4
T_0 (°C)	930†	1,200	875
T_m (°C)	825	750	715
T_d (°C)	855	800	760
λ_d	2.7‡	1.8‡	2.05‡
μ_d (Pas)	$10^{5.47}$ – $10^{6.65}$ §	$10^{4.86}$ – $10^{5.45}$ §	$10^{5.91}$ – $10^{6.95}$ §
H (m)	1,500–3,500	4,500	3,000
w_0 (m)	0.1	0.1	0.1
Calculated			
T_x (°C)	828.4	817.3	734.6
Independent estimates			
t_{RT} (days)	(10,500), 1,400–30	(330), 75–60	–
$1/\lambda_m$	$10^{3.64}$ – $10^{1.44}$ ($10^{4.9}$)	$10^{1.63}$ – $10^{0.67}$ ($10^{2.1}$)	–

The symbols are defined in Table 1. The references from which values are taken are given in the Supplementary Tables.

* Ranges were used to calculate $1/\lambda_m$ and average values in parentheses were used as input in the unzipping model.

† Value lowered from the original 1,050 °C so that $T_x > T_m$.

‡ Average value of three rheological models (Supplementary Discussion).

§ Ranges given by three rheological models (Supplementary Discussion). We used the melt viscosity and/or water content given by the literature (Supplementary Tables).

|| Values in parentheses are extremes that delimit the outer edges of the natural domains (Fig. 3).

Ranges given by three rheological models (Supplementary Discussion). $1/\lambda_m$ was calculated by adding 5 vol.% crystals to the mush.

METHODS SUMMARY

The mush is assumed to be initially motionless and isothermal⁹. The intrusion is assumed to pond at the base of the mush and to interact with it by heat transfer only. Such under-accretion stems from defining a mush as a crystal-rich magma that does not react in a brittle fashion to the deformation rates of active magmatic processes. This leaves nearly solidified magma bodies (>80 vol.% crystals) out of our analysis because they are subject to brittle penetration and over-accretion. We do not need to take into account the thermal history of the intrusion because remobilizing a semi-rigid magma body can easily be triggered by a modest amount of fresh magma ponding beneath it (Supplementary Discussion).

At the beginning of the reheating, a layer forms by conductive melting of the overlying mush until it reaches a critical thickness at which convection starts within the now-mobile layer. To a first-order approximation, this remobilization can be adequately described as a homogeneous fluid with stair-step rheology (Supplementary figures). The mobile layer then grows at a faster rate such that the heat transferred through the convecting layer balances that needed to melt the overlying mush. We solved this classical moving-boundary problem¹⁴ analytically to express the layer growth rate \dot{d} as a function of time. Under the combined effects of buoyancy and interface perturbations due to convection²⁴, the thickening mobile layer is prone to Rayleigh–Taylor instabilities. We formulated the evolution of an instability of amplitude w starting on top of the convecting layer. Overturn starts when the growth rate of the large-scale instabilities \dot{w} is faster than that of the stable front²⁴ \dot{d} . The partial overturn not only exports homogenized material from the

mobile layer to the top of the system but also causes enough mixing within the mush to bring together crystals that were far apart²⁵.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 2 June 2010; accepted 7 January 2011.

Published online 2 March 2011.

- Bachmann, O. & Bergantz, G. W. Deciphering magma chamber dynamics from styles of compositional zoning in large silicic ash flow sheets. *Rev. Mineral. Geochem.* **69**, 651–674 (2008).
- Charlier, B. L. A., Bachmann, O., Davidson, J. P., Dungan, M. A. & Morgan, D. J. The upper crustal evolution of a large silicic magma body: evidence from crystal-scale Rb–Sr isotopic heterogeneities in the Fish Canyon Magmatic System, Colorado. *J. Petrol.* **48**, 1875–1894 (2007).
- Davidson, J. P., Hora, J. M., Garrison, J. M. & Dungan, M. A. Crustal forensics in arc magmas. *J. Volcanol. Geotherm. Res.* **140**, 157–170 (2005).
- Miller, C. F. *et al.* Growth of plutons by incremental emplacement of sheets in crystal-rich host: evidence from Miocene intrusions of the Colorado river region, Nevada, USA. *Tectonophysics* (in the press); doi:10.1016/j.tecto.2009.07.011.
- Boyce, J. W. & Hervig, R. L. Magmatic degassing histories from apatite volatile stratigraphy. *Geology* **36**, 63–66 (2008).
- Miller, J. S., Matzel, J. E. P., Miller, C. F., Burgess, S. D. & Miller, R. B. Zircon growth and recycling during the assembly of large, composite arc plutons. *J. Volcanol. Geotherm. Res.* **167**, 282–299 (2007).
- Wallace, P. J. Volatiles in subduction zone magmas: concentrations and fluxes based on melt inclusion and volcanic gas data. *J. Volcanol. Geotherm. Res.* **140**, 217–240 (2005).
- Hildreth, W. Volcanological perspectives on Long Valley, Mammoth Mountain, and Mono Craters: several contiguous but discrete systems. *J. Volcanol. Geotherm. Res.* **136**, 169–198 (2004).
- Huber, C., Bachmann, O. & Manga, M. Homogenization processes in silicic magma chambers by stirring and mushification (latent heat buffering). *Earth Planet. Sci. Lett.* **283**, 38–47 (2009).
- Bachmann, O. & Bergantz, G. W. Gas percolation in upper-crustal magma bodies as a mechanism for upward heat advection and rejuvenation of near-solidus magma bodies. *J. Volcanol. Geotherm. Res.* **149**, 85–102 (2006).
- Couch, S., Sparks, R. S. J. & Carroll, M. R. Mineral disequilibrium in lavas explained by convective self-mixing in open magma chambers. *Nature* **411**, 1037–1039 (2001).
- Huber, C., Bachmann, O. & Dufek, J. The limitations of melting on the reactivation of silicic mushes. *J. Volcanol. Geotherm. Res.* **195**, 97–105 (2010).
- Mahood, G. A. Second reply to comment of R.S.J. Sparks, H.E. Huppert and C.J.N. Wilson on “Evidence for long residence times of rhyolitic magma in the Long Valley magmatic system: the isotopic record in the precaldera lavas of Glass Mountain”. *Earth Planet. Sci. Lett.* **99**, 395–399 (1990).
- Huppert, H. E. & Sparks, R. S. J. The generation of granitic magmas by intrusion of basalt into continental crust. *J. Petrol.* **29**, 599–624 (1988).
- Scailliet, B., Whittington, A., Martel, C., Pichavant, M. & Holtz, F. Phase equilibrium constraints on the viscosity of silicic magma II: implications for mafic-silicic mixing processes. *Trans. R. Soc. Edinb. Earth Sci.* **91**, 61–72 (2000).
- Bachmann, O. & Bergantz, G. W. Rejuvenation of the Fish Canyon magma body: a window into the evolution of large-volume silicic magma systems. *Geology* **31**, 789–792 (2003).
- Annen, C. & Sparks, R. S. J. Effects of repetitive emplacement of basaltic intrusions on thermal evolution and melt generation in the crust. *Earth Planet. Sci. Lett.* **203**, 937–955 (2002).
- Michaut, C. & Jaupart, C. Ultra-rapid formation of large volumes of evolved magma. *Earth Planet. Sci. Lett.* **250**, 38–52 (2006).
- Wiebe, R. A. & Collins, W. J. Depositional features and stratigraphic sections in granitic plutons: implications for the emplacement and crystallization of granitic magma. *J. Struct. Geol.* **20**, 1273–1289 (1998).
- de Silva, S. & Gosnold, W. D. Episodic construction of batholiths: insights from the spatiotemporal development of an ignimbrite flare-up. *J. Volcanol. Geotherm. Res.* **167**, 320–335 (2007).
- Self, S., Zhao, J.-X., Holasek, R. E., Torres, R. C. & King, A. J. in *Fire and Mud; Eruptions and Lahars of Mount Pinatubo, Philippines* (eds Newhall, C. G. & Punongbayan, R. S.) 1089–1115 (University of Washington Press, 1996).
- Devine, J. D., Rutherford, M. J., Norton, G. E. & Young, S. R. Magma storage region processes inferred from geochemistry of Fe–Ti oxides in andesitic magma, Soufrière Hills volcano, Montserrat, W.I. *J. Petrol.* **44**, 1375–1400 (2003).
- Bachmann, O., Dungan, M. A. & Lipman, P. W. The Fish Canyon magma body, San Juan volcanic field, Colorado: rejuvenation and eruption of an upper-crustal batholith. *J. Petrol.* **43**, 1469–1503 (2002).
- Ke, Y. & Solomatov, V. S. Plume formation in strongly temperature-dependent viscosity fluids over a very hot surface. *Phys. Fluids* **16**, 1059–1063 (2004).
- Ruprecht, P., Bergantz, G. W. & Dufek, J. Modeling of gas-driven magmatic overturn: tracking of phenocryst dispersal and gathering during magma mixing. *Geochim. Geophys. Geosyst.* **9**, Q07017, doi:10.1029/2008GC002022 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank O. Bachmann and C. Huber for discussions. The work was funded partially by the ERC grant 202844 under the European FP7.

Author Contributions A.B. and G.W.B. participated to the theoretical construction of the model. A.B. realized the numerical implementation and produced the first draft of the paper, which both authors then discussed.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.B. (burgisse@cnrs-orleans.fr).

METHODS

The mush is assumed to be initially motionless and isothermal⁹. The intrusion is assumed to pond at the base of the mush and to interact with it by heat transfer only. Such under-accretion stems from defining a mush as a crystal-rich magma that does not react in a brittle fashion to the deformation rates of active magmatic processes. This leaves nearly solidified magma bodies (>80 vol.% crystals) out of our analysis because they are subject to brittle penetration and over-accretion. We do not need to take into account the thermal history of the intrusion because remobilizing a semi-rigid magma body can easily be triggered by a modest amount of fresh magma ponding beneath it (Supplementary Discussion).

At the beginning of the reheating, the mobile layer melts conductively and the temperature at the base of the layer T_b is the average of intrusion temperature T_0 and mush temperature T_m (see Supplementary figures for a detailed geometry and Supplementary Tables for a full symbol list). In a standard procedure for such moving-boundary problems¹⁴, we first treat the general case of the growth of a convecting mobile layer before applying the result to conductive growth. The Rayleigh number in the mobile layer is defined by the properties at the average mid-layer temperature T_d :

$$Ra = \frac{\Delta \rho g d^3}{\kappa \mu_d} \quad (1)$$

where d is the layer thickness, g is the acceleration due to gravity, κ is the thermal diffusivity, μ_d is the layer viscosity, and $\Delta \rho = \varepsilon \alpha \rho_0 (T_b - T_d)$ is the density contrast between the hottest and the coldest parts of the layer (ε is the melt volume fraction, α is the thermal expansion coefficient, and ρ_0 is the reference density). The mid-layer temperature can be determined given the expected viscosity variation across the mobile layer λ_d , expressed as the ratio of the highest and lowest viscosity values²⁶:

$$T_d = T_\lambda + \frac{T_b - T_\lambda}{1 + \lambda_d^{-1/6}} \quad (2)$$

In our case it will be close to half the temperature difference between the top and the base of the layer. The mobile layer is growing over time by slowly melting the mush (that is, bringing the mush from T_m to the temperature at which the magma becomes mobile, T_λ):

$$\dot{d} = \frac{F}{\rho_0 (c_p (T_\lambda - T_m) + f_m L_m)} \quad (3)$$

where c_p is the heat capacity of the mush, f_m is the weight fraction of mush that melts, L_m is the mush latent heat, and F is the heat flux. By definition, the Nusselt number in the mobile layer once convection takes place is:

$$Nu = \frac{F d}{k (T_d - T_\lambda)} \quad (4)$$

where k is the thermal conductivity. Within the layer, the influence of convection on heat transfer can be assessed by relating Nu to Ra :

$$Nu = a Ra^b \quad (5)$$

The coefficients a and b have been determined either empirically^{26,27} or theoretically²⁸ (Supplementary Tables). Coefficients proposed from experimental work on convection of variable viscosity fluids yield quasi-identical results. Scale analysis, on the other hand, suggests that a regime change occurs in the middle of the parameter range relevant to magmatic convection, thus framing the experimental laws but creating a discontinuity that is cumbersome to handle for the simple model we develop here. We thus used the experimentally based values given by ref. 27.

The heat flux can be replaced in equation (3) by its expression using equations (4) and (5):

$$\dot{d} = A d^{3b-1} \quad (6)$$

where

$$A = \frac{a \kappa c_p (T_d - T_\lambda)}{c_p (T_\lambda - T_m) + f_m L_m} \left[\frac{\Delta \rho g}{\kappa \mu_d} \right]^b$$

Here, the remobilization process (that is, bringing the mush from T_m to the mobile layer temperature T_d) has been assumed to have negligible effects on the values of c_p , k and κ , thus allowing us to use single values for the mush and the mobile layer. Taking into account that b falls between 1/5 and 1/3 (Supplementary Tables) and that $\dot{d}(t=0) = 0$, integration of equation (6) gives:

$$t = \frac{d^{2-3b}}{A(2-3b)} \quad (7)$$

The critical Rayleigh number, $Ra_c = 1,708$, can be used to obtain the critical thickness of the mobile layer d_c :

$$d_c = \left(\frac{\kappa Ra_c \mu_d}{\Delta \rho g} \right)^{1/3} \quad (8)$$

The time at which convection starts t_c can be obtained by setting $a = 1$ and $b = 0$ so as to have a Nusselt number of one, and evaluating equation (7) at $d = d_c$. The mush becomes entirely remobilized when the mobile layer fills the entire chamber, which occurs at the time t_H given by evaluating equation (7) at $d = H$.

The growth rate of the mobile layer can now be expressed as:

$$\dot{d} = A^{1/(2-3b)} [(2-3b)t]^{1/(2-3b)-1} \quad (9)$$

Assuming that under-plating occurs over an area of H^2 (that is, a cubic mush reservoir with an intrusion spreading beneath its entire floor), the total thermal energy T_J needed for remobilization is given by:

$$T_J = H^2 \int_0^t F(\phi) d\phi \quad (10)$$

where ϕ is an integration variable for time. Using equations (3) and (9) to express F and integrating the result yields:

$$T_J = H^2 \rho_0 c_p (T_\lambda - T_m) [(2-3b)t]^{1/(2-3b)} \quad (11)$$

Dividing T_J by $H^2 \rho_b \left[c_{p(b)} \left(\frac{T_0 - T_m}{2} \right) + f_b L_b \right]$ yields the minimum total thickness of the basalt layer, H_b . (f_b is the weight fraction of basalt that crystallizes when cooling from T_0 to T_b , ρ_b is the basalt density, $c_{p(b)}$ is the basalt heat capacity and L_b is the latent heat of crystallization). The maximum thickness of the basalt layer can be estimated using a simple conductive approach by which the temperature decrease at the interface between intrusion and mush is given by the complementary error function²⁹ $\text{erfc}(H_b / \sqrt{\kappa_b t}) (T_0 - T_m) / 2$ (where κ_b is the basalt thermal diffusivity). Our approach assumes that this difference, say 0.5 °C, remains small over t_{RT} , and the above equation can be used to calculate H_b .

Unzipping starts when the growth rate of the large-scale instabilities is faster than the growth rate of the mobile layer itself. The duration from intrusion emplacement to unzipping is t_{RT} . Canright and Morris³⁰ described the growth of a perturbation of amplitude $w(t)$ at the interface between two fluid layers of contrasted viscosities. They show that the perturbation growth law depends on the rheology of the fluids, either Newtonian or non-Newtonian. Below, we adapt their resolution in order to obtain salient laws for $w(t)$ and estimate t_{RT} by solving $\dot{d} = \dot{w}$.

Newtonian rheology. A Newtonian mush has a linear relationship between strain rate $\dot{\gamma}$ and shear stress τ :

$$\tau = \mu_m \dot{\gamma} \quad (12)$$

The development of an instability starting after the onset of convection in the mobile layer follows³⁰:

$$w(t) = w_0 \exp \left(\frac{\Delta \rho_m g d}{\mu_m} \tilde{\sigma} (t - t_c) \right) \quad (13)$$

where $\Delta \rho_m$ is the density contrast between the mush and the mobile layer, w_0 is the initial amplitude of the instability, and $\tilde{\sigma}$ is the dimensionless growth rate. The density variation in the mobile layer is a combination of the change in crystal content and the reheating of the interstitial liquid: $\Delta \rho_m = f_m (\rho_c - \rho_0) + \varepsilon \alpha \rho_0 (T_d - T_m)$, where ρ_c is the average density of the solid phases that melt. To reduce the degrees of freedom of the model, we fixed $\rho_c = 2,700 \text{ kg m}^{-3}$, which corresponds to plagioclase, a phase generally abundant in mushes. The dimensionless growth rate is given by³⁰:

$$\tilde{\sigma} = \frac{\lambda_m (s/K - 1)(C - 1) + (K - \beta)(c - 1)}{2\lambda_m (Cc - 1 + \beta K^2) + \lambda_m^2 (S + \beta K)(s - K) + (S - \beta K)(s + K)} \quad (14)$$

where the symbols $s = \sinh(K)$, $c = \cosh(K)$, $S = \sinh(\beta K)$, $C = \cosh(\beta K)$, $\beta = d/(H - d)$ and $K = 4\pi(H - d)/\eta$ have been used. $\lambda_m = \mu_d/\mu_m$ is the viscosity ratio between mobile layer and mush and η is the wavelength of the instability. In the parameter range of interest ($\beta < 10^{-0.2}$, $10^{-8} < \lambda_m < 10^{-0.6}$), there is always a wavelength η_{\max} for which the dimensionless growth rate $\tilde{\sigma}_{\max}$ is maximum. It can be found by solving:

$$\frac{d\tilde{\sigma}}{dK} = 0 \quad (15)$$

The instability onset time t_{RT} can then be obtained by solving $\dot{d} = \dot{w}$:

$$[(2-3b)At_{RT}]^{\frac{1}{2-3b}} = w_0 \exp\left(\frac{\Delta\rho_m g}{\mu_m} d\tilde{\sigma}_{\max} t_{RT}\right) \quad (16)$$

$$\frac{\Delta\rho_m g}{\mu_m} \left(d\tilde{\sigma}_{\max} + d\dot{\tilde{\sigma}}_{\max} t_{RT} + d\ddot{\tilde{\sigma}}_{\max} t_{RT}\right)$$

Non-Newtonian rheology

The rheologic law of a non-Newtonian mush can be expressed as a power law of exponent $M > 1$ and the consistency coefficient $\mu_{m(0)}$:

$$\tau = \mu_{m(0)} \dot{\gamma}^{1/M} \quad (17)$$

The growth of a perturbation of amplitude $w(t)$ becomes³⁰:

$$w(t) = w_0 \left[1 - (M-1) \left(\frac{\Delta\rho_m g}{4\mu_{m(0)}} \right)^M w_0^{M-1} d(t-t_c) \right]^{\frac{1}{1-M}} \quad (18)$$

The time taken for the large-scale instability to start growing faster than the mobile layer can thus be evaluated by deriving \dot{w} from equation (18) and substituting it into the right-hand side of equation (16):

$$[(2-3b)At_{RT}]^{\frac{1}{2-3b}} = d \left(\frac{w(t_{RT})\Delta\rho_m g}{4\mu_{m(0)}} \right)^M \quad (19)$$

Numerical resolution. For the Newtonian case, the time until the onset of large-scale instabilities, t_{RT} , is found by fixing all parameters ($\Delta\rho_m$, μ_d , μ_m , w_0) and successively solving equations (15) and (16) numerically. Setting $\tilde{\sigma} = 0$ in equation (16) causes errors in t_{RT} of $<9\%$, and setting in addition $\tilde{d} = 0$ in the right-hand side of equation (16) increases these errors to $<16\%$. For the non-Newtonian case, t_{RT} is found by fixing all parameters and solving equation (19) numerically (Supplementary figures). All equations are solved using the Newton–Raphson root-finding algorithm. The thickness of the mobile layer at which unzipping starts d_{RT} can then be found using equation (7). The time t_p for the large-scale Rayleigh–Taylor instability to reach the roof of the reservoir can be calculated, if its

amplitude is small, by solving equation (13) for $w(t_p) = H - d$ with $\tilde{\sigma}_{\max}$ evaluated at d_{RT} .

Estimating the strain rate and the shear stress applied to the mush during unzipping is useful to determine which rheology applies best to the unzipping process. To a first-order approximation, the strain rate can be obtained by evaluating the stress caused by the growth of the instability near its tip. For a Newtonian rheology, it is independent of time:

$$\dot{\gamma} = \frac{dw}{dw} = \frac{\Delta\rho_m g \tilde{\sigma} d_{RT}}{\mu_d} \quad (20)$$

For a non-Newtonian rheology, the strain rate depends on time and is evaluated at the onset of unzipping:

$$\dot{\gamma} = \frac{dw}{dw} = w(t_{RT})^{M-1} M d_{RT} \left(\frac{\Delta\rho_m g}{4\mu_{m(0)}} \right)^M \quad (21)$$

The shear stress can then be evaluated by using the rheological laws (equations (12) and (17)), respectively. If the mush rheology is such that a yield stress exists, the mush will be set in motion when the static stress applied to the mush by the buoyancy of mobile layer reaches a certain value. This stress is a function of the roughness of the melting interface, which we relate here to the size of the initial perturbation:

$$\tau_{\text{static}} = \Delta\rho_m g w_0 \quad (22)$$

26. Schaeffer, N. & Manga, M. Interaction of rising and sinking mantle plumes. *Geophys. Res. Lett.* **28**, 455–458 (2001).
27. Manga, M. & Weeraratne, D. Experimental study of non-Boussinesq Rayleigh–Bernard convection at high Rayleigh and Prandtl numbers. *Phys. Fluids* **11**, 2969–2976 (1999).
28. Grossmann, S. & Lohse, D. Thermal convection for large Prandtl numbers. *Phys. Rev. Lett.* **86**, 3316–3319 (2001).
29. Crank, J. *The Mathematics of Diffusion* (Oxford University Press, 1975).
30. Canright, D. & Morris, S. Buoyant instability of a viscous film over a passive fluid. *J. Fluid Mech.* **255**, 349–372 (1993).

Human-specific loss of regulatory DNA and the evolution of human-specific traits

Cory Y. McLean^{1*}, Philip L. Reno^{2,3*†}, Alex A. Pollen^{2*}, Abraham I. Bassan², Terence D. Capellini², Catherine Guenther^{2,3}, Vahan B. Indjeian^{2,3}, Xinhong Lim², Douglas B. Menke^{2,3†}, Bruce T. Schaar², Aaron M. Wenger¹, Gill Bejerano^{1,2} & David M. Kingsley^{2,3}

Humans differ from other animals in many aspects of anatomy, physiology, and behaviour; however, the genotypic basis of most human-specific traits remains unknown¹. Recent whole-genome comparisons have made it possible to identify genes with elevated rates of amino acid change or divergent expression in humans, and non-coding sequences with accelerated base pair changes^{2–5}. Regulatory alterations may be particularly likely to produce phenotypic effects while preserving viability, and are known to underlie interesting evolutionary differences in other species^{6–8}. Here we identify molecular events particularly likely to produce significant regulatory changes in humans: complete deletion of sequences otherwise highly conserved between chimpanzees and other mammals. We confirm 510 such deletions in humans, which fall almost exclusively in non-coding regions and are enriched near genes involved in steroid hormone signalling and neural function. One deletion removes a sensory vibrissae and penile spine enhancer from the human androgen receptor (*AR*) gene, a molecular change correlated with anatomical loss of androgen-dependent sensory vibrissae and penile spines in the human lineage^{9,10}. Another deletion removes a forebrain subventricular zone enhancer near the tumour suppressor gene growth arrest and DNA-damage-inducible, gamma (*GADD45G*)^{11,12}, a loss correlated with expansion of specific brain regions in humans. Deletions of tissue-specific enhancers may thus accompany both loss and gain traits in the human lineage, and provide specific examples of the kinds of regulatory alterations^{6–8} and inactivation events¹³ long proposed to have an important role in human evolutionary divergence.

To discover human-specific deletions (hDELs) on a genome-wide scale (Fig. 1a), we identified regions of the chimpanzee genome¹⁴ with a clear macaque orthologue, but no closely related sequence in humans (Supplementary Information). This identified 37,251 ancestral primate sequences lost in humans, spanning 34.0 megabases (Mb; 1.17%) of the chimpanzee genome. To find deletions most likely to produce functional consequences, we also identified highly conserved sequences spanning 70.0 Mb (2.41%) of the chimpanzee genome, producing a conservative estimate of pan-mammalian sequence under purifying selection in chimpanzee¹⁴. The intersection of these deletion and conservation surveys identified 583 regions with high sequence conservation that are surprisingly deleted in humans, which we term hCONDELs (Supplementary Table 2).

Of the 583 predicted hCONDELs, 510 (87.5%) were validated independently by single human sequence reads spanning both sides of the deletions (Supplementary Information). Of the sequence-validated hCONDELs, 85% seemed fixed in human trace archives. Experimental amplifications across human diversity panels confirmed human deletion of 39/39 randomly chosen hCONDELs, and showed fixation of 31/32 hCONDELs also fixed in the trace archives (Supplementary

Tables 3–5). Of the sequence-validated hCONDELs, 88% are missing from the draft Neanderthal genome¹⁵ (Supplementary Information), in agreement with the length of the common lineage between humans and Neanderthals following the divergence with chimpanzee.

The 583 hCONDELs cover 3.96 Mb (0.14%) of the chimpanzee genome, remove an average of 95 base pairs (bp) of conserved sequence, and are found on every nuclear chromosome except Y, where macaque sequence is unavailable (Fig. 1b). hCONDELs have a median size of 2,804 bp, and show a skew towards G/C-poor regions compared to chimpanzee genome-wide averages (Fig. 1c and Supplementary Fig. 2). We find no evidence for enrichment in areas with high recombination ($P > 0.9$) or pericentromeric and subtelomeric regions ($P > 0.5$) (Supplementary Information). Thus, loss of sequences is unlikely to be a secondary consequence of higher mutation rates in such regions, a possibility that has been debated for the accelerated sequence changes seen in other human non-coding elements^{4,5,16}.

Only 1 of 510 sequence-validated hCONDELs removes a protein-coding region, corresponding to a previously known 92 bp deletion in *CMAH* (ref. 17). All other sequence-validated hCONDELs map to non-protein coding regions of the genome (355 intergenic and 154 intronic). Similar highly conserved non-coding sequences frequently correspond to regulatory elements controlling expression of nearby genes¹⁸. Whereas hCONDELs cannot be tested directly for evidence of positive human selection, we can examine whether hCONDELs are preferentially located near genes with particular functions, using the Genomic Regions Enrichment of Annotations Tool (GREAT)¹⁹. We performed enrichment analysis using simulations to explicitly account for the overrepresentation of highly conserved elements near particular classes of genes (Supplementary Information). This analysis indicates that hCONDELs are significantly enriched near genes involved in steroid hormone receptor signalling and neural function (Table 1), and near genes encoding fibronectin-type-III- or CD80-like immunoglobulin C2-set domains (Supplementary Table 10).

To compare sequences lost in other lineages, we used a similar computational approach to identify conserved sequences lost specifically in chimpanzee or mouse (termed cCONDELs and mCONDELs, respectively). We identified 344 cCONDELs and 350 mCONDELs validated by sequence reads spanning predicted deletions (Supplementary Information). cCONDELs and mCONDELs show enrichment for synapse and glutamate receptors (Supplementary Table 11) and metal ion binding (Supplementary Table 12), respectively, but not for the categories found near hCONDELs.

To explore further the possible functions of hCONDELs near steroid hormone signalling genes, we examined a 60.7 kilobase (kb) human deletion flanking the androgen receptor (*AR*) locus (Fig. 2a) in detail. *AR* is required for response of tissues to circulating androgens, and thus represents a strong candidate locus for characteristic changes of

¹Department of Computer Science, Stanford University, Stanford, California 94305, USA. ²Department of Developmental Biology, Stanford University School of Medicine, Stanford, California 94305, USA. ³Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305, USA. [†]Present addresses: Department of Anthropology, Pennsylvania State University, University Park, Pennsylvania 16801, USA (P.L.R.); Department of Genetics, University of Georgia, Athens, Georgia 30602, USA (D.B.M.).

*These authors contributed equally to this work.

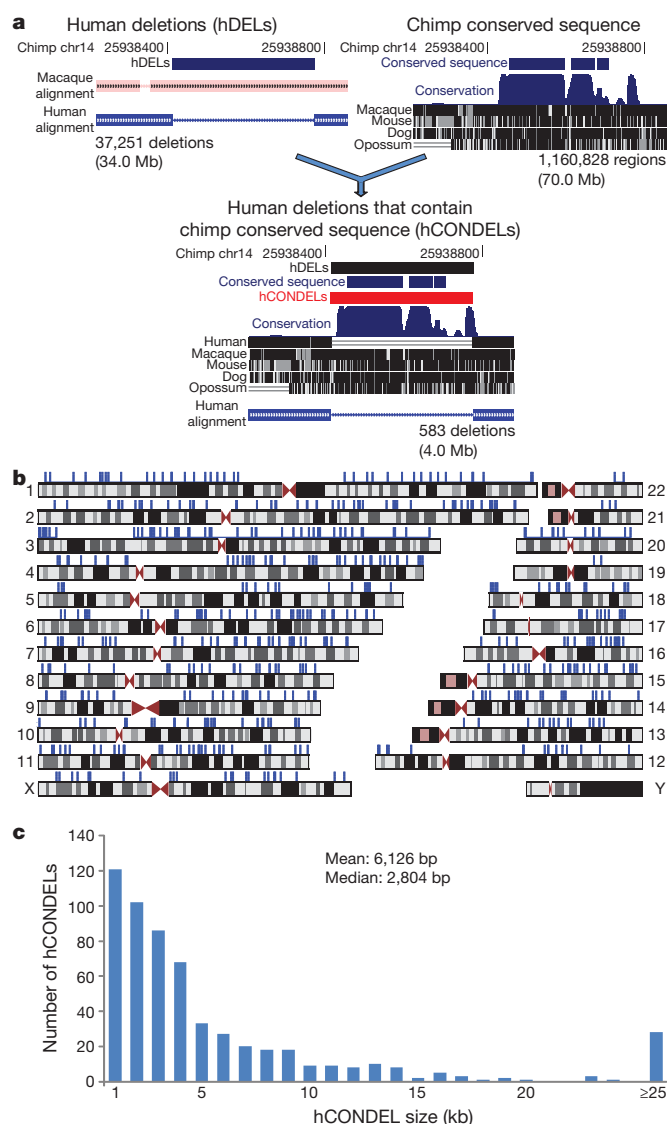


Figure 1 | Hundreds of sequences highly conserved between chimpanzee and other species are deleted in humans. **a**, Computational approach used to discover human-specific deletions of functional DNA: identification of ancestral chimpanzee genomic sequences deleted in human; discovery of chimpanzee genomic sequences highly conserved in other species; and detection of human-specific deletions that remove one or more chimpanzee conserved sequences. Total chimpanzee sequence identified in each step is displayed beneath each graphic. **b**, Human genomic locations of the 583 hCONDELs. hCONDELs are displayed as blue ticks above the many locations where they are missing. **c**, Size distribution of hCONDELs.

secondary sexual traits known in the human lineage^{20,21}. Within this human-specific deletion lies a ~5 kb region containing highly conserved non-coding sequences. We cloned the corresponding 4,839 bp chimpanzee and 7,654 bp mouse regions and tested their capacity to drive expression of an hsp68 (also known as Hspa1b) basal promoter-*lacZ* reporter gene during normal mouse development. Chimpanzee and mouse constructs both drove consistent *lacZ* expression in the facial vibrissae and genital tubercle of five or more independent transgenic embryos (Fig. 2b–g), and the mouse sequence also drove expression in hair follicles.

lacZ expression was specifically located in the mesoderm surrounding vibrissae follicles, and in the superficial mesoderm within the presumptive glans of the developing genital tubercle (Fig. 2d,e), a site of known AR expression (Fig. 2h). Four stable mouse lines all showed expression in the superficial tissue underlying epidermal spines of the penis (Fig. 2i). Previous studies have shown that AR is expressed in mesenchyme surrounding developing epithelial structures²² and is required for normal development of vibrissae and penile spines (see below). These results indicate that the human deletion removes a conserved enhancer sequence that directs expression in a spatially-restricted subset of the AR expression pattern. The chimpanzee sequence also drives significant reporter gene expression when tested in human foreskin fibroblasts, indicating that upstream pathways regulating the enhancer are still intact in humans (chimpanzee activity 2.2–5.9-fold greater than human deletion/basal control vector, $P < 10^{-7}$).

Interestingly, humans show obvious morphological differences at the anatomical locations controlled by the enhancer. Sensory vibrissae develop in many mammals including chimpanzees, macaques and mice⁹. In contrast, humans lack sensory vibrissae⁹. Vibrissae development is clearly androgen-responsive, as castration shortens vibrissae in mice, and excess testosterone increases growth²³ (Fig. 2j).

Profound changes have also evolved in the genitalia of humans compared to other animals. Many mammals have keratinized epidermal spines overlying tactile receptors in the glans dermis^{10,20}. Penile spine growth is androgen-dependent, as primates lose spines upon castration, and treatment with exogenous testosterone restores spine formation²⁴ (Fig. 2k). Mice with AR protein-coding mutations fail to form penile spines²⁵, confirming an essential role for AR in penile spine development. Our results show that humans have lost an ancestral penile spine enhancer from the AR locus. Humans also fail to form the penile spines commonly found in other animals, including chimpanzees, macaques and mice^{10,20,25} (Fig. 2l). Simplified penile morphology tends to be associated with monogamous reproductive strategies in primates²⁰. Ablation of spines decreases tactile sensitivity and increases the duration of intromission²⁰, indicating their loss in the human lineage may be associated with the longer duration of copulation in our species relative to chimpanzees²⁰. This fits with an adaptive suite, including feminization of the male canine dentition, moderately-sized testes with low sperm motility, and concealed ovulation with permanently enlarged mammary glands²⁰, that suggests our ancestors evolved numerous morphological characteristics associated with pair-bonding and increased paternal care²¹.

Table 1 | Summary of annotation enrichments of hCONDELs

Ontology: Term	Expected number of hCONDELs	Observed number of hCONDELs	Fold enrichment	Binomial <i>P</i> -value ¹⁹
GO Molecular Function: Steroid hormone receptor activity	4.7	14	2.96	3.7×10^{-4}
Entrez Gene: Neural genes	141.3	180	1.27	1.1×10^{-4}
MGI Expression in Theiler Stage 21:				
Hindbrain	49.9	79	1.58	3.4×10^{-5}
Cerebral cortex	42.1	68	1.62	7.0×10^{-5}
Brain, ventricular layer	29.9	52	1.74	9.4×10^{-5}
Midbrain	30.5	52	1.70	1.6×10^{-4}
InterPro Protein Domains:				
Fibronectin, type III	16.7	34	2.03	1.0×10^{-4}
CD80-like, immunoglobulin C2-set	2.1	8	3.84	1.4×10^{-3}

Showing only non-redundant terms that are enriched after accounting for both multiple testing and for the tendency of conserved elements to be found near particular classes of genes (Supplementary Information).

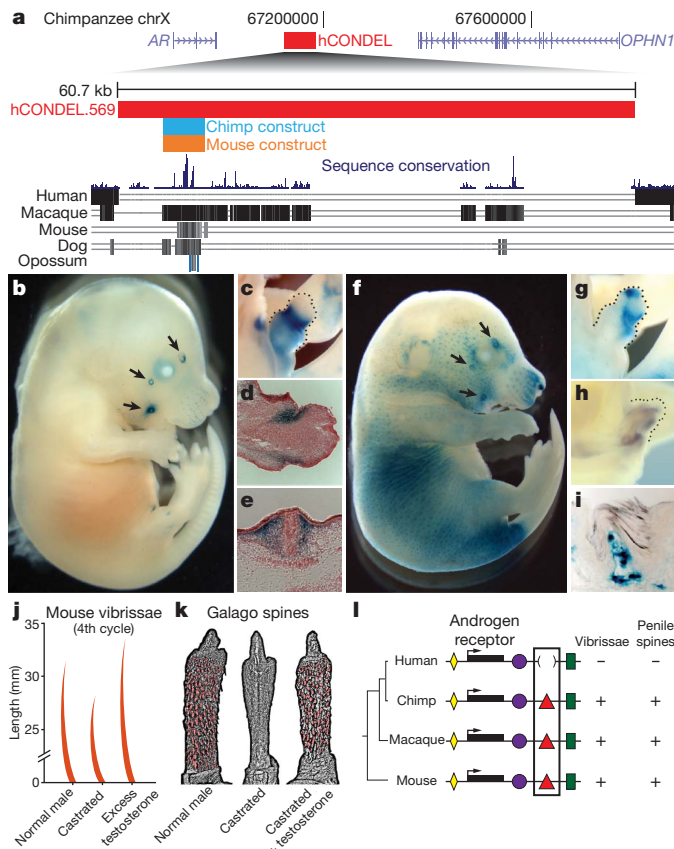


Figure 2 | Transgenic analysis of a chimpanzee and mouse AR enhancer region missing in humans. **a**, Top panel: 1.1 Mb region of the chimpanzee X chromosome. The red bar shows the position of a 60.7-kb human deletion removing a well-conserved chimpanzee enhancer between the *AR* and *OPHN1* genes. Bottom panel: multiple species comparison of the deleted region, showing sequences aligned between chimpanzee and other mammals. Blue and orange bars represent chimpanzee and mouse sequences tested for enhancer activity in transgenic mice. The chimpanzee sequence drives *lacZ* expression in **b**, facial vibrissae (arrows), and **c**, genital tubercle (dotted line) of E16.5 mouse embryos. Histological sections reveal strongest staining in superficial mesenchyme of **d**, the prospective glans of the genital tubercle, and **e**, dermis surrounding the base of sensory vibrissae. The mouse enhancer also drives consistent expression in **f**, facial vibrissae, **g**, genital tubercle, and hair follicles of E16.5 embryos. **h**, Endogenous *AR* is expressed in the genital tubercle (dotted line) as demonstrated by *in situ* hybridization. **i**, Histological section of a 60-day-old transgenic mouse penis showing postnatal *lacZ* expression in dermis of penile spines. Vibrissae and penile spines are androgen-dependent, as shown by **j**, changes in vibrissae length in castrated and testosterone-treated mice and **k**, loss and recovery of penile spines of a castrated and testosterone-treated primate (*Galago crassicaudatus*) (modified from refs 23 and 24). **l**, Model depicting multiple conserved tissue-specific enhancers (coloured shapes) surrounding *AR* coding sequences (black bars) of different species. Loss of an ancestral vibrissae/penile spine enhancer in humans is correlated with corresponding loss of sensory vibrissae and penile spines.

Sensory vibrissae and penile spines are examples of morphological structures lost in humans. However, deletion of enhancers could also be associated with tissue expansion, by removing regulatory sequences from genes that control developmental patterns of cell proliferation, death or migration. One of the most dramatic tissue expansions in human evolution is increased size of the cerebral cortex¹. Using a curated resource of gene expression domains during mouse development, we find that hCONDELs are significantly enriched near genes expressed during cortical neurogenesis ($P = 7 \times 10^{-5}$, Table 1). Within this set, hCONDELs are preferentially located near genes acting as suppressors of cell proliferation or migration ($P = 0.003$) (Supplementary Information).

To test the function of one such hCONDEL, we analysed further a deletion removing a 3,181-bp region located next to the tumour suppressor gene growth arrest and DNA-damage-inducible, gamma (*GADD45G*) (Fig. 3a). This hCONDEL removes a forebrain-specific p300 binding site¹⁸, which predicts enhancer sequences with strong specificity. The chimpanzee version of this sequence, and a smaller 546 bp mouse sequence overlapping the p300 enhancer-binding region, both drove *lacZ* expression in the developing ventral telencephalon and diencephalon in at least five independent embryonic day 14.5 (E14.5) transgenic embryos (Fig. 3b–i), confirming that the ancestral sequence corresponds to a conserved forebrain-specific enhancer. The chimpanzee sequence also drives significant gene expression when tested in immortalized human fetal neural progenitor cells, suggesting the enhancer would also affect transcription if still present in humans (1.7–2-fold greater expression than human deletion/basal control vector, $P < 10^{-7}$).

Histological sections of transgenic embryos show that the chimpanzee and mouse sequences drive *lacZ* expression in the subventricular zone (SVZ) of the septum, the preoptic area, and in regions of the ventral thalamus and hypothalamus (Fig. 3c–e, g–i). The *lacZ* expression matches a sub-domain of the expression pattern of *GADD45G*, which is normally expressed throughout the telencephalon SVZ and diencephalon²⁶. The preoptic area generates inhibitory interneurons that migrate to the neocortex and ventral telencephalon²⁷. The ventral thalamus domain corresponds to a region that generates inhibitory interneurons, which have notably increased in proportion in the human thalamus^{28,29}.

Expression in brain subventricular zones is particularly interesting in light of previous suggestions that increased proliferation of SVZ

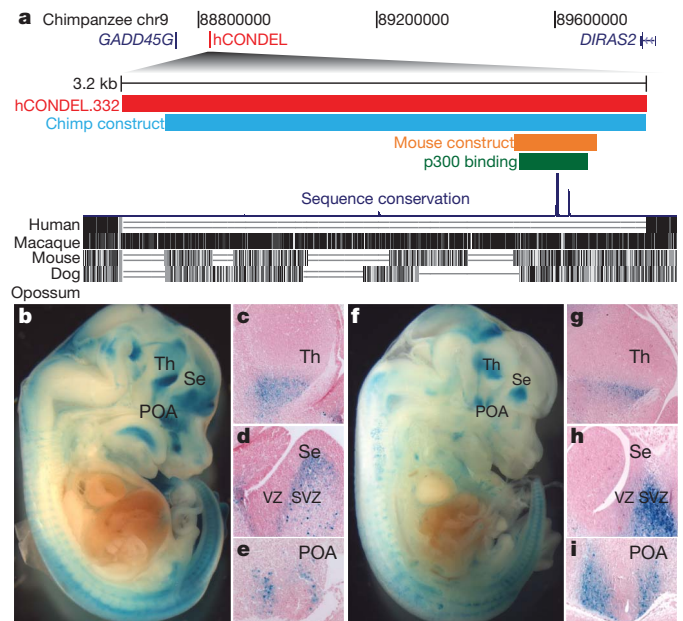


Figure 3 | Transgenic analysis of a chimpanzee and mouse forebrain enhancer missing from a tumour suppressor gene in humans. **a**, Top panel: 1.3 Mb region of the chimpanzee chromosome 9. The red bar illustrates a 3,181 bp human-specific deletion removing a conserved chimpanzee enhancer located downstream of *GADD45G*. Bottom panel: multiple species comparison of the deleted region, showing sequences aligned between chimpanzee and other mammals. The green bar represents a mouse forebrain-specific p300 binding site¹⁸, and the blue and orange bars represent chimpanzee and mouse sequences tested for enhancer activity in transgenic mice. The chimpanzee (**b–e**) and mouse sequence (**f–i**) both drive consistent *lacZ* expression in E14.5 mouse embryos in the ventral thalamus (**c**, **g**), the SVZ of the septum (**d**, **h**), and the preoptic area (**e**, **i**). Increased production of neuronal subtypes from these regions may contribute to thalamic and cortical expansion in humans^{27–30}. All sections are sagittal with anterior to right. POA, preoptic area; Se, septum; SVZ, subventricular zone; Th, thalamus; VZ, ventricular zone.

intermediate progenitors underlies the evolutionary expansion of the neocortex in primates³⁰. *GADD45G* normally represses cell cycle and can activate apoptosis, and somatic loss of *GADD45G* expression is clearly linked with excess tissue growth in human pituitary adenomas^{11,12}. Our results show that species-specific loss of an ancestral SVZ enhancer has clearly occurred in the human lineage, providing a plausible molecular basis for increasing production of particular neuronal cell types by regulatory changes in a tumour suppressor gene.

We cannot exclude the possibility that loss of *AR* and *GADD45G* enhancers has occurred because of relaxed selection following other genetic changes that have led to anatomical differences in the human lineage. However, based on the previously established role of *AR* in vibrissae and penile spine development, and of *GADD45G* in negative regulation of tissue proliferation, we think it probable that deletions of tissue-specific enhancers in these genes have contributed to both loss and expansion of particular tissues during human evolution. The full set of hCONDELs may contain loci associated with other human-specific characteristics, a possibility that can now be tested by further functional studies of these conserved non-coding sequences that are surprisingly missing from the human genome.

METHODS SUMMARY

We examined 2,696 Mb (92.67%) of the chimpanzee genome aligning with single regions in humans, removing assembly and alignment artefacts, segmental duplications, and other regions with complex genome histories (Supplementary Information). Chimpanzee sequences aligning to macaque but not human were identified as human-specific deletions. We identified 70.0 Mb (2.41%) of the chimpanzee genome as highly conserved by calculating the fraction of identical base pairs within multiple sequence alignments using a series of sliding window criteria (Supplementary Table 1). hCONDELs were validated by searching for individual human sequence reads in the NCBI Trace Archives spanning predicted deletions; and by testing 39 hCONDELs by experimental amplification from individuals of 23 human populations (Coriell Cell Repositories, Supplementary Information). Gene enrichments were analysed using GREAT (great.stanford.edu)¹⁹. For each ontology we ran 1,000 simulations over 510 size-matched random deletions overlapping one or more chimpanzee conserved elements to derive a multiple test *P*-value threshold for enrichment. Functional enhancer assays were carried out by injecting chimp and mouse expression constructs into FVB embryos (Xenogen Biosciences and Cyagen Biosciences) and staining for lacZ expression activity at different developmental stages, or by transfecting chimp and mouse enhancer constructs into human foreskin fibroblasts (System Biosciences) or ReNcell CX Human Neural Progenitor cells (Chemicon), and comparing expression to human-deletion/control constructs using Galacto-Light Plus (Applied Biosciences) (Supplementary Information).

Received 14 June; accepted 22 December 2010.

- Varki, A. & Altheide, T. Comparing the human and chimpanzee genomes: searching for needles in a haystack. *Genome Res.* **15**, 1746–1758 (2005).
- Bustamante, C. D. *et al.* Natural selection on protein-coding genes in the human genome. *Nature* **437**, 1153–1157 (2005).
- Khaitovich, P. *et al.* Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309**, 1850–1854 (2005).
- Pollard, K. S. *et al.* An RNA gene expressed during cortical development evolved rapidly in humans. *Nature* **443**, 167–172 (2006).
- Prabhakar, S. *et al.* Human-specific gain of function in a developmental enhancer. *Science* **321**, 1346–1350 (2008).
- King, M. C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116 (1975).
- Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008).
- Chan, Y. F. *et al.* Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **327**, 302–305 (2010).

- Muchlinski, M. N. A comparative analysis of vibrissa count and infraorbital foramen area in primates and other mammals. *J. Hum. Evol.* **58**, 447–473 (2010).
- Hill, W. C. O. Note on the male external genitalia of the chimpanzee. *Proc. Zool. Soc. Lond.* **116**, 129–132 (1946).
- Zerbini, L. F. *et al.* NF- κ B-mediated repression of growth arrest- and DNA-damage-inducible proteins 45 α and γ is essential for cancer cell survival. *Proc. Natl Acad. Sci. USA* **101**, 13618–13623 (2004).
- Zhang, X. *et al.* Loss of expression of *GADD45 γ* , a growth inhibitory gene, in human pituitary adenomas: implications for tumorigenesis. *J. Clin. Endocrinol. Metab.* **87**, 1262–1267 (2002).
- Olson, M. When less is more: gene loss as an engine of evolutionary change. *Am. J. Hum. Genet.* **64**, 18–23 (1999).
- The Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87 (2005).
- Green, R. E. *et al.* A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).
- Duret, L. & Galtier, N. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu. Rev. Genomics Hum. Genet.* **10**, 285–311 (2009).
- Chou, H. H. *et al.* A mutation in human CMP-sialic acid hydroxylase occurred after the *Homo-Pan* divergence. *Proc. Natl Acad. Sci. USA* **95**, 11751–11756 (1998).
- Visel, A. *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858 (2009).
- McLean, C. Y. *et al.* GREAT improves functional interpretation of *cis*-regulatory regions. *Nature Biotechnol.* **28**, 495–501 (2010).
- Dixon, A. F. *Primate Sexuality* (Oxford University Press, 1998).
- Lovejoy, C. O. Reexamining human origins in light of *Ardipithecus ramidus*. *Science* **326**, 74e1–74e8 (2009).
- Crocoll, A., Zhu, C. C., Cato, A. C. & Blum, M. Expression of androgen receptor mRNA during mouse embryogenesis. *Mech. Dev.* **72**, 175–178 (1998).
- Ibrahim, L. & Wright, E. A. Effect of castration and testosterone propionate on mouse vibrissae. *Br. J. Dermatol.* **108**, 321–326 (1983).
- Dixon, A. F. Effects of testosterone on the sternal cutaneous glands and genitalia of the male greater galago (*Galago crassicaudatus crassicaudatus*). *Folia Primatol. (Basel)* **26**, 207–213 (1976).
- Murakami, R. A histological study of the development of the penis of wild-type and androgen-insensitive mice. *J. Anat.* **153**, 223–231 (1987).
- Gohlke, J. M. *et al.* Characterization of the proneural gene regulatory network during mouse telencephalon development. *BMC Biol.* **6**, 15 (2008).
- Gelman, D. M. *et al.* The embryonic preoptic area is a novel source of cortical GABAergic interneurons. *J. Neurosci.* **29**, 9380–9389 (2009).
- Vue, T. Y. *et al.* Characterization of progenitor domains in the developing mouse thalamus. *J. Comp. Neurol.* **505**, 73–91 (2007).
- Arcelli, P., Frasson, C., Regondi, M. C., De Biasi, S. & Spreafico, R. GABAergic neurons in mammalian thalamus: a marker of thalamic complexity? *Brain Res. Bull.* **42**, 27–37 (1997).
- Kriegstein, A., Noctor, S. & Martínez-Cerdeño, V. Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion. *Nature Rev. Neurosci.* **7**, 883–890 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. DeGusta for providing chimpanzee DNA, S. McConnell and P. Buckmaster for discussions, and M. Hiller and C. Barr for ontology analysis support. This work was supported in part by a Bio-X graduate fellowship (C.Y.M.), a Ruth L. Kirschstein NRSA post-doctoral fellowship (1 F32 HD062137-01, P.L.R.), a National Defense Science and Engineering Graduate fellowship (A.A.P.), a National Science Scholarship of the Agency of Science, Technology, and Research, Singapore (X.L.), a Stanford Graduate Fellowship (A.M.W.), an Edward Mallinckrodt, Jr. Foundation grant (G.B.), and National Institute of Health grants R01 HD059862 (G.B.), R01 HG005058 (G.B.) and P50 HG002568 (D.M.K.). G.B. is a Packard Fellow, Searle Scholar, Microsoft Faculty Fellow and an Alfred P. Sloan Fellow. D.M.K. is an investigator of the Howard Hughes Medical Institute.

Author Contributions G.B. and D.M.K. conceived the investigation; C.Y.M. performed the computational analyses; P.L.R., A.A.P., A.I.B., T.D.C., C.G., V.B.I., X.L., D.B.M. and B.T.S. performed the experiments; C.Y.M., P.L.R., A.A.P., B.T.S., A.M.W., G.B. and D.M.K. analysed the data; and C.Y.M., P.L.R., A.A.P., G.B. and D.M.K. wrote the paper with contributions from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to G.B. (bejerano@stanford.edu) or D.M.K. (kingsley@stanford.edu).

Co-adjuvant effects of retinoic acid and IL-15 induce inflammatory immunity to dietary antigens

R. W. DePaolo^{1*}, V. Abadie^{1*}, F. Tang¹, H. Fehlner-Peach¹, J. A. Hall^{2,3}, W. Wang¹, E. V. Marietta^{4,5}, D. D. Kasarda⁶, T. A. Waldmann⁷, J. A. Murray⁸, C. Semrad¹, S. S. Kupfer¹, Y. Belkaid², S. Guandalini⁹ & B. Jabri^{1,9,10}

Under physiological conditions the gut-associated lymphoid tissues not only prevent the induction of a local inflammatory immune response, but also induce systemic tolerance to fed antigens^{1,2}. A notable exception is coeliac disease, where genetically susceptible individuals expressing human leukocyte antigen (HLA) HLA-DQ2 or HLA-DQ8 molecules develop inflammatory T-cell and antibody responses against dietary gluten, a protein present in wheat³. The mechanisms underlying this dysregulated mucosal immune response to a soluble antigen have not been identified. Retinoic acid, a metabolite of vitamin A, has been shown to have a critical role in the induction of intestinal regulatory responses^{4–6}. Here we find in mice that in conjunction with IL-15, a cytokine greatly upregulated in the gut of coeliac disease patients^{3,7}, retinoic acid rapidly activates dendritic cells to induce JNK (also known as MAPK8) phosphorylation and release the proinflammatory cytokines IL-12p70 and IL-23. As a result, in a stressed intestinal environment, retinoic acid acted as an adjuvant that promoted rather than prevented inflammatory cellular and humoral responses to fed antigen. Altogether, these findings reveal an unexpected role for retinoic acid and IL-15 in the abrogation of tolerance to dietary antigens.

Induction of regulatory intestinal responses to oral antigens prevents the subsequent development of systemic T-helper type-1 (T_H1) responses to those antigens, a phenomenon referred to as oral tolerance². An exception to this is coeliac disease, where patients mount a T_H1 immune response to dietary gluten³. To determine whether the presence of IL-15 may affect intestinal homeostasis, we first examined its effects on the generation of inducible regulatory T cells (T_{reg}) expressing the transcription factor forkhead box P3 (Foxp3). Under physiological conditions, mesenteric lymph node (MLN) dendritic cells drive *de novo* differentiation of T_{reg} cells^{4–6}. In the presence of IL-15-stimulated MLN dendritic cells, T_{reg}-cell generation from unfractionated CD4⁺ T cells (Fig. 1a) or naive CD44^{lo} CD4⁺ T cells (Supplementary Fig. 1a) was impaired. Further, IL-15 had no effect on T_{reg}-cell differentiation in the presence of dendritic cells lacking the IL-2–IL-15-receptor-β/γ-chain (IL-2–IL-15-Rβ/γc) heterodimeric signalling receptor complex⁸ (Supplementary Fig. 2a) and in dendritic-cell-free systems (Supplementary Figs 1b, 5), demonstrating that IL-15 was acting at the level of dendritic cells and not T cells to block T_{reg} generation. To assess the relevance of our *in vitro* findings, we tested the response to fed chicken ovalbumin (OVA), a model antigen used in oral tolerance experiments, in D^d-IL-15 transgenic mice⁹ that over-express IL-15 in the lamina propria and MLN but not in the intestinal epithelium (Supplementary Fig. 3). In agreement with our *in vitro* observations, the number of naive OT-II RAG1^{−/−} CD4⁺ T cells converted into T_{reg} cells was significantly reduced in OVA-fed D^d-IL-15 transgenic mice compared to wild-type mice (Fig. 1b). Retinoic acid

(RA), which is reported to have a critical role in T_{reg} differentiation^{4–6}, further decreased the conversion of T_{reg} cells *in vivo* in OVA-fed D^d-IL-15 transgenic mice (Fig. 1c).

To determine the mechanisms by which IL-15-stimulated dendritic cells prevent the conversion of T_{reg} cells and to assess further the role of RA, we used splenic dendritic cells (Supplementary Fig. 2b), which, unlike MLN dendritic cells, lack the ability to produce constitutively high levels of RA^{4,6}. Conditioned media obtained from IL-15-treated splenic dendritic cells contained high levels of IL-12p70 and IL-23, but no IL-6 (Supplementary Fig. 4a), and decreased T_{reg}-cell conversion (Supplementary Fig. 2c). The proinflammatory properties of IL-15 were enhanced further in the presence of RA (Supplementary Fig. 4a). As previously reported¹⁰, IL-12p70 significantly suppressed T_{reg} differentiation at concentrations found in IL-15-conditioned dendritic cell supernatant (Supplementary Fig. 4b), and this effect was enhanced by IL-23 in a dose-dependent manner (Supplementary Fig. 4b). The requirement for IL-12 in IL-15-mediated suppression of T_{reg} differentiation was demonstrated using IL-12p40-deficient dendritic cells (Fig. 1d and Supplementary Fig. 4c) and neutralizing anti-IL-12p40 antibody (Supplementary Fig. 4d).

Inhibition of T_{reg}-cell conversion in the presence of IL-15 was accompanied by the induction of T_H1 responses *in vitro* (Fig. 2a and Supplementary Fig. 6a) and *in vivo* in OVA-fed D^d-IL-15 transgenic mice (Fig. 2b and Supplementary Fig. 6c). As expected, the ability of IL-15 to induce T_H1 responses *in vitro* (Fig. 2c) and *in vivo* (Fig. 2f) was dependent on IL-12. RA further promoted (Fig. 2a, b and Supplementary Fig. 6a–c) and was central for T_H1 polarization in the presence of IL-15 (Fig. 2d, e). In addition, RA acted in concert with IL-12p70 at the level of T cells and with IL-6 to amplify T_H1-cell differentiation (Supplementary Figs 5, 6b) and T_H17-cell responses (Supplementary Fig. 7a–d), respectively. The critical role of IL-15 in the induction of inflammatory T-cell responses to dietary antigen was demonstrated by blocking IL-15 signalling *in vivo* using either a neutralizing anti-IL-15 or IL-15/IL-2Rβ antibody¹¹ (Fig. 2f, g and Supplementary Figs 6d, 7e, f). Altogether these observations indicate a sequential model whereby IL-15 first acts in concert with RA to induce IL-12 and IL-23 in MLN dendritic cells (Fig. 2d and data not shown). Along with RA, these inflammatory mediators then operate at the level of T cells to promote T_H1-cell differentiation and, when IL-6 is present¹², T_H17-cell differentiation. Although our finding of the adjuvant effect of RA is unexpected within the field of mucosal immunity, it is consistent with its usage as a beneficial proinflammatory adjuvant in anti-tumour immunity^{13,14}.

On the basis of our observations indicating that IL-15 and RA act primarily at the level of dendritic cells to disrupt intestinal immune homeostasis, we investigated which signalling pathway was critical for

¹Department of Medicine, University of Chicago, Chicago, Illinois 60637, USA. ²Mucosal Immunology Unit, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. ³Immunology Graduate Group, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. ⁴Department of Dermatology, Mayo Clinic College of Medicine, Rochester, Minnesota 55905, USA. ⁵Department of Immunology, Mayo Clinic College of Medicine, Rochester, Minnesota 55905, USA. ⁶US Department of Agriculture, Agricultural Research Service, Western Regional Research Center, 800 Buchanan Street, Albany, California 94710, USA. ⁷Metabolism Branch, National Cancer Institute, Bethesda, Maryland 20892-1374, USA. ⁸Department of Medicine, Division of Gastroenterology and Hepatology, Mayo Clinic College of Medicine, Rochester, Minnesota 55905, USA. ⁹Department of Pediatrics, University of Chicago, Chicago, Illinois 60637, USA. ¹⁰Department of Pathology, University of Chicago, Chicago, Illinois 60637, USA.

*These authors contributed equally to this work.

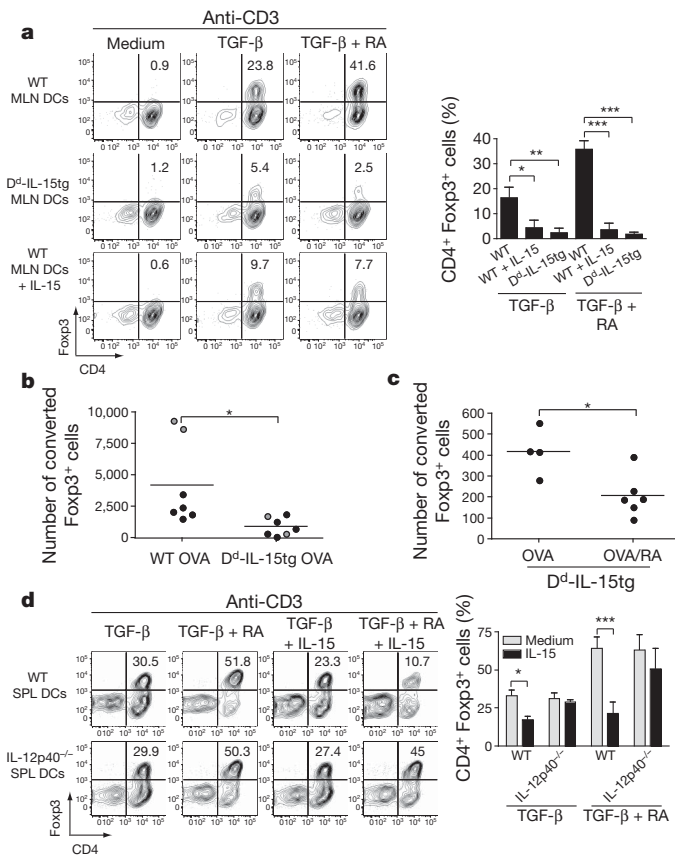


Figure 1 | IL-15-activated dendritic cells in the presence of retinoic acid prevent induction of Foxp3⁺ regulatory T cells. **a**, 10^5 CD4⁺ Foxp3[−] T cells were cultured with 4×10^4 MLN dendritic cells (DCs) isolated from wild-type (WT) or D^d-IL-15 transgenic (D^d-IL-15tg) mice with anti-CD3 alone or combined with IL-15, TGF- β and RA. The percentages of Foxp3⁺ cells are shown. Graph depicts pooled data \pm s.e.m. ($n = 3$). **b**, RAG1^{−/−} OT-II CD45 congenic CD25[−] CD4⁺ T cells were transferred into wild-type and D^d-IL-15 transgenic mice that were fed OVA in drinking water for five days (black dots) or by gavage every other day for 10 days (grey dots). T_{reg}-cell conversion was assessed in the MLN by intracellular staining for Foxp3 and detected by flow cytometry. The absolute numbers of converted CD4⁺ Foxp3⁺ T cells are shown. Data are representative of two experiments performed independently. The decrease in the number of converted T_{reg} cells was associated with a significant decrease in the number of detectable transferred T cells in D^d-IL-15 transgenic mice (data not shown). This is probably due to the inability to detect inflammatory T cells that are more susceptible to cell death than Foxp3⁺ T_{reg} cells, which express anti-apoptotic factors. **c**, Ly5.2⁺ OT-II T cells were transferred into Ly5.1⁺ and D^d-IL-15tg-Ly5.1⁺ recipient mice that were fed OVA or OVA and RA five times during ten days. The absolute number of CD4⁺ Foxp3⁺ Ly5.2⁺ converted T cells in the MLN is shown as in **a**. **d**, As in **a**, CD4⁺ Foxp3[−] T cells were cultured with splenic (SPL) dendritic cells isolated from wild-type and IL-12p40^{−/−} mice. The percentages of Foxp3⁺ cells are indicated. Graph depicts three pooled experiments \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (unpaired Student's *t*-test).

their proinflammatory effects. IL-15 rapidly induced JNK phosphorylation in a dose-dependent manner in dendritic cells (Fig. 3a). The ability of RA to synergize with IL-15 to promote IL-12p70 (Fig. 3b) and IL-23 (Supplementary Fig. 8a) production by dendritic cells and significantly reduce T_{reg} differentiation (Fig. 3c) was paralleled by its ability to synergize with IL-15 to promote JNK phosphorylation (Fig. 3d). The co-adjuvant effects of RA were mediated via the RA receptor α (RAR α)/JNK signalling pathway (Fig. 3e–g and Supplementary Fig. 8b–d) and may involve previously reported¹⁵ non-genomic signalling effects of the RAR (Fig. 3h). We do not exclude the possibility that RA may signal *in vivo* via an alternative RAR such as RAR β , however. Of note, extracellular signal-regulated kinase

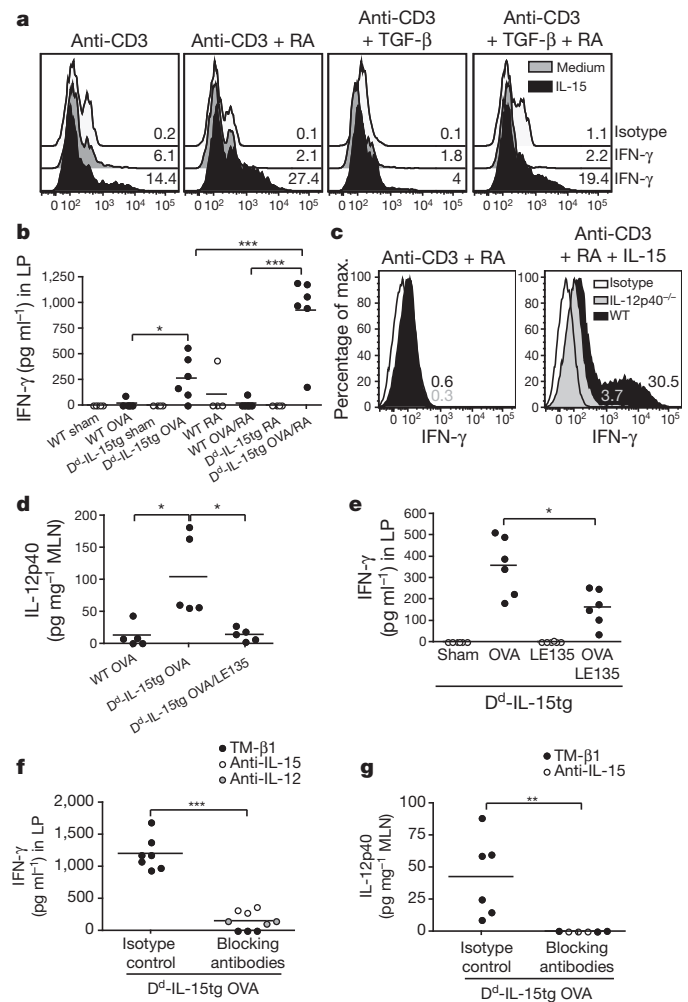


Figure 2 | Retinoic acid exerts an adjuvant effect on IL-15-mediated inflammatory T-cell responses. **a**, CD4⁺ T cells were cultured with wild-type splenic dendritic cells with the indicated cytokines. Representative histograms gated on CD4⁺ T cells show IFN- γ expression ($n = 5$). **b**, D^d-IL-15 transgenic and wild-type mice were fed PBS (sham), OVA, RA, or a mixture of OVA and RA. IFN- γ secretion by lamina propria (LP) cells re-stimulated for 24 h with OVA. The results are the means of triplicate samples obtained from two independent experiments. **c**, CD4⁺ T cells were cultured with splenic dendritic cells isolated from wild-type or IL-12p40^{−/−} mice as described in **a**. Intracellular staining for IFN- γ of gated CD4 T cells is shown. Results are representative of two experiments. **d**, Levels of IL-12p40 in the MLN of wild-type and D^d-IL-15 transgenic mice fed OVA, or a mixture of OVA and the RAR antagonist LE135. The results are the means of triplicate samples obtained from two independent experiments. Similar results were obtained for IL-12p70 and IL-23 (data not shown). **e**, IFN- γ secretion by lamina propria cells isolated from D^d-IL-15 transgenic mice fed PBS (sham), OVA and LE135. The results are the means of triplicate samples obtained from two independent experiments. **f**, **g**, D^d-IL-15 transgenic mice were fed OVA and treated with blocking anti-IL-12p40, anti-IL-15 and TM β -1 (anti-IL-2R β) or isotype control monoclonal antibodies. The levels of IL-12p40 in the MLN (**f**) and IFN- γ in lamina propria cells re-stimulated overnight with OVA (**g**) were quantified. When anti-IL-15 and anti-IL-12 treatment experiments were performed in parallel, control mice received a mixture of corresponding isotype controls. Data represent two pooled experiments ($n = 6$ mice per group) except for the anti-IL-12 treatment ($n = 3$ individual mice). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (unpaired Student's *t*-test).

(ERK; also known as MAPK1) and p38 MAPK (also known as MAPK14) were dispensable (Supplementary Fig. 8e).

An intriguing aspect of coeliac disease pathogenesis is how an inflammatory T_H1 response is induced against dietary gluten proteins. Owing to the large number of proline residues they contain, gluten

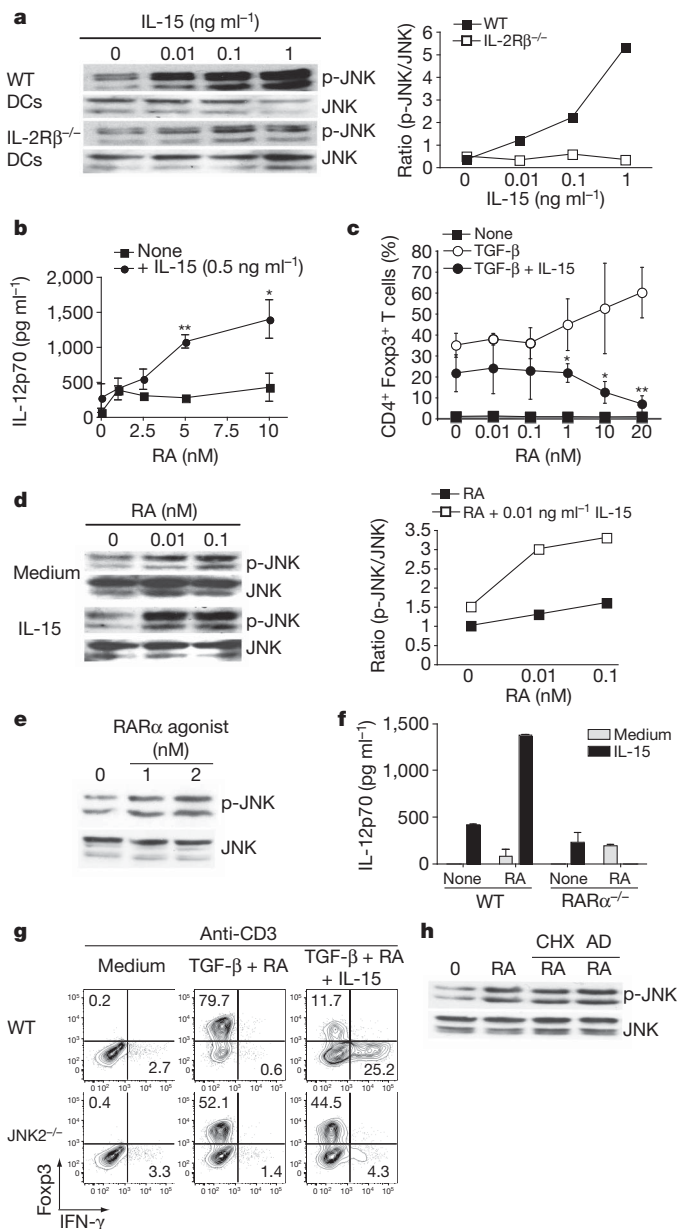


Figure 3 | Retinoic acid and IL-15 act in synergy to induce dendritic cells with proinflammatory properties in a JNK-dependent manner.

a, Concentration-dependent JNK phosphorylation in wild-type or IL-2Rβ^{-/-} bone-marrow-derived dendritic cells upon IL-15 stimulation analysed by western blot (left panel) and quantified (right panel). **b**, IL-12p70 secretion after overnight culture of wild-type splenic dendritic cells with increasing doses of RA, with and without IL-15. Results are mean values ± s.e.m. (*n* = 3). **c**, CD4⁺ Foxp3⁺ T cells were cultured with wild-type splenic dendritic cells with anti-CD3 alone or combined with IL-15, TGF-β and increasing doses of RA. The percentages of Foxp3⁺ cells are shown. Graph depicts pooled data ± s.e.m. (*n* = 3). **d**, Concentration-dependent JNK phosphorylation in wild-type bone-marrow-derived dendritic cells upon stimulation with IL-15 (0.01 ng ml⁻¹) and increasing doses of RA stimulation by western blot (left panel) and quantified (right panel). **e**, Concentration-dependent JNK phosphorylation in wild-type bone-marrow-derived dendritic cells upon stimulation with a RARα agonist (AM580). **f**, IL-12p70 secretion after overnight culture of wild-type and RARα^{-/-} bone-marrow-derived dendritic cells with IL-15 alone or combined with RA. Data are shown as means and s.e.m. (*n* = 2). **g**, CD4⁺ Foxp3⁺ T cells were cultured with splenic dendritic cells isolated from wild-type or JNK2^{-/-} mice with the indicated cytokines. The percentages of Foxp3⁺ and IFN-γ⁺ among CD4⁺ T cells are indicated. **h**, JNK phosphorylation in bone-marrow-derived dendritic cells pretreated with cyclohexamide (CHX) or actinomycin D (AD) before stimulation with 0.1 nM RA. Results are representative of two independent experiments. **P* < 0.05, ***P* < 0.01 (unpaired Student's *t*-test).

proteins are resistant to enzymatic degradation, leading to the generation of long peptides that are selectively presented in the gut by HLA-DQ2 or HLA-DQ8 molecules³. However, these observations alone fail to explain why 40% of the population expresses the coeliac-disease-associated HLA-DQ2 and HLA-DQ8 molecules, yet induction of gluten-specific inflammatory CD4⁺ and CD8⁺ T-cell responses^{16,17} occurs in less than 2% of these individuals³. In addition, patients with coeliac disease typically develop antibodies to gluten and to the enzyme tissue transglutaminase 2 (TG2; also known as TGM2)^{3,18}. To test the hypothesis that IL-15 overexpressed in the lamina propria may break tolerance to dietary gluten, D^d-IL-15 transgenic mice that have levels of IL-15 in the lamina propria comparable to those observed in the lamina propria of coeliac disease patients⁷ (compare Supplementary Fig. 3a and Fig. 4e) and lack IL-15 upregulation in intestinal epithelial cells (IECs) (Supplementary Fig. 3b), were crossed onto humanized HLA-DQ8 transgenic mice¹⁹ (DQ8-D^d-IL-15 transgenic mice). The presence of IFN-γ-producing anti-gliadin T cells (Fig. 4a and Supplementary Figs 9a, 10a, b, d, e) in conjunction with anti-gliadin (Fig. 4b) and anti-TG2 (Fig. 4c) antibodies and intraepithelial lymphocytosis (Fig. 4d) in DQ8-D^d-IL-15 transgenic mice fed gliadin, phenocopies an early stage of coeliac disease²⁰ where villous atrophy is absent. The normal intestinal morphology and intraepithelial lymphocytes (IELs) phenotype in D^d-IL-15 transgenic mice (data not shown) mirrors human observations indicating that IL-15 expression by IECs drives IELs to become fully licensed killer cells and induce villous atrophy^{3,21}. In accordance with human studies^{22,23}, anti-gliadin T_H17 cells were detected only at a very low frequency (Supplementary Figs 9b, c, 10c). Similar to the observations in OVA-fed D^d-IL-15 transgenic mice, the induction of inflammatory anti-gliadin T_H1 and T_H17 responses was dependent on (Supplementary Fig. 11) and promoted by RA (Fig. 4a and Supplementary Figs 9a–c, 10). Of note, gliadin-fed DQ8 transgenic mice had a slight induction in IFN-γ-producing T cells, which may be related to the reported innate effects of gliadin³. Supporting the hypothesis that IL-15 may disrupt tolerance to gluten in coeliac disease patients by inducing IL-12, the levels of IL-15 and IL-12p70 were correlated in the lamina propria of active coeliac disease patients (Fig. 4e). This discovery also gives a functional foundation to the identification of IL-12A as a genetic risk factor for coeliac disease by genome-wide association studies²⁴.

Our study reveals that in the presence of IL-15, RA has unforeseen co-adjuvant properties that induce T_H1 immunity to fed antigens (Fig. 4f). It indicates further that under infectious conditions associated with induction of IL-15 and IL-6 in the intestinal mucosa, RA will also promote T_H17 immunity. These observations caution against the use of vitamin A and RA for the treatment of autoimmunity and inflammatory intestinal disorders associated with high levels of IL-15. Indeed, a causal relationship between retinoids used for the treatment of acne and inflammatory bowel disease has been implicated in a subset of patients²⁵. Conversely, these findings provide an explanation as to why children suffering from vitamin A deficiency in developing countries²⁶ respond less efficiently to oral vaccines than children from developed countries^{26,27}, and also indicate that engineering mucosal vaccines that induce IL-15 may be beneficial owing to their ability to induce concomitantly protective IgA antibodies and T_H1 immunity.

More generally, our study supports the concept that there are no 'unconditional' suppressive factors, and that integration of tissue and exogenous signals determine the class of the immune response, which ultimately needs to be tailored to the tissue and the antigen. In line with the idea that the same proinflammatory factors trigger different immunological outcomes depending on the tissue in which they are induced, we found that the ability of IL-12 to inhibit T_{reg}-cell induction was blocked by butyrate, a metabolite produced by commensal bacteria present in the colon but not in the small bowel (data not shown).

One final aspect of our study is that we may have in hand a long-awaited physiopathologically relevant murine model mimicking the early stages of coeliac disease. This model is unique in that development

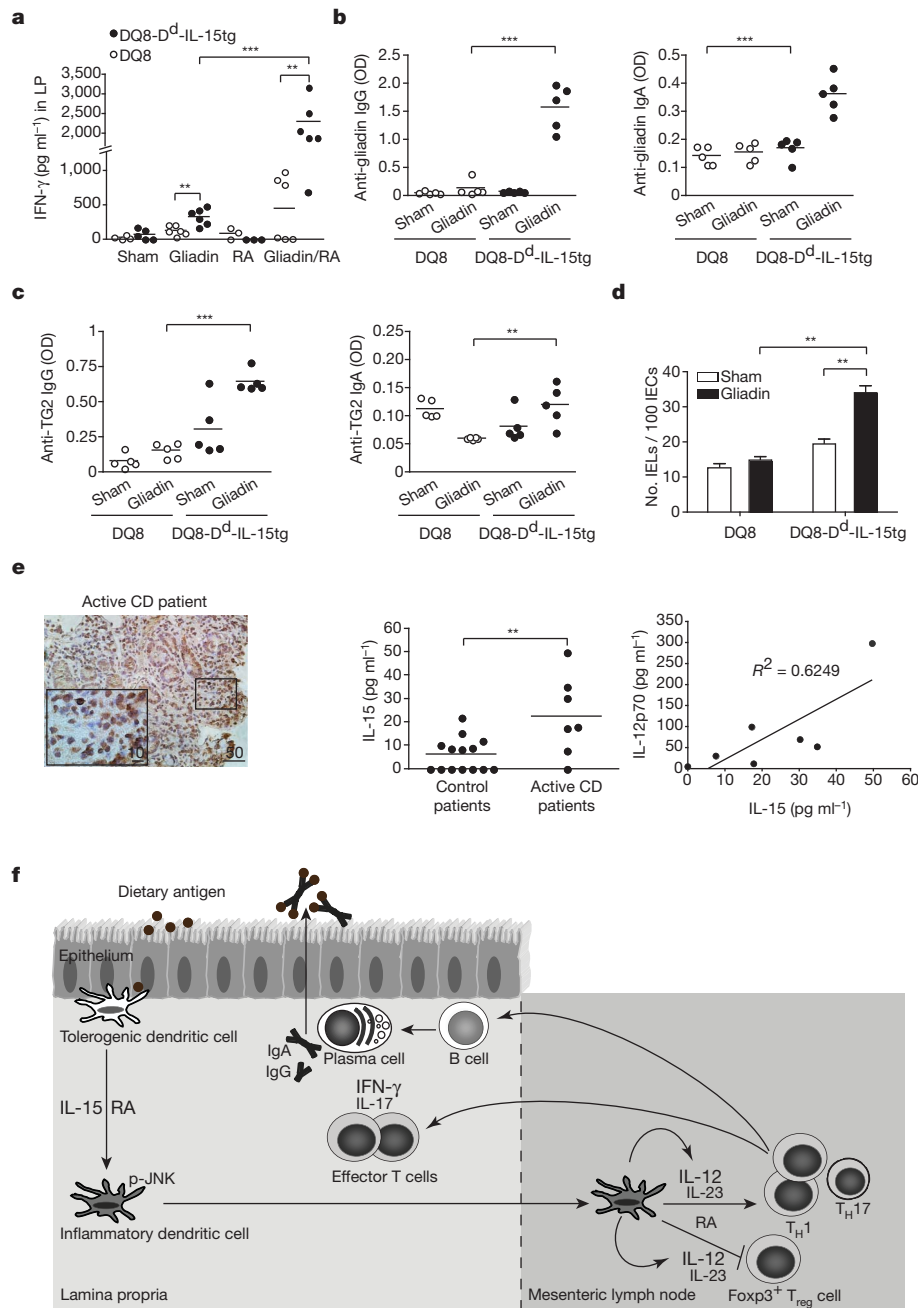


Figure 4 | DQ8-D^d-IL-15 transgenic mice fed gliadin mimic early stages of coeliac disease reflecting dysregulation in the adaptive immune response to gluten. **a–d**, DQ8 and DQ8-D^d-IL-15 transgenic mice were fed gliadin every other day for ten days. **a**, IFN- γ secretion by lamina propria cells after overnight culture with gliadin. **b**, **c**, Anti-gliadin IgG, anti-gliadin IgA, anti-TG2 IgG and anti-TG2 IgA titres from serum collected fifteen days after feeding. Antibody titres were detected by ELISA and calculated according to the formula: (OD_{450 nm} of sample – OD_{450 nm} of blank) \times serum dilution, where OD is optical density. **d**, Quantification of IELs among IECs in small intestines fifteen days after the last gliadin feeding. Two independent counts of IELs among 200 IECs were performed for each mouse. Data are shown as means and s.e.m. ($n = 3$ individual mice). Similar results were obtained in another set of experiments where mice were fed with α -gliadin. **e**, IL-15 and IL-12 expression in the lamina propria of coeliac disease patients. Immunohistochemical

stainings for IL-15 in gut tissue from an active coeliac disease (CD) patient (left panel). Lamina propria cells were harvested from biopsies obtained from control ($n = 14$) or active coeliac disease patients ($n = 7$) and assayed for levels of IL-15 and IL-12p70 by ELISA (middle and right panels). Equal concentration of total proteins was analysed for each sample. $^{**}P < 0.01$, $^{***}P < 0.001$ (unpaired Student's t -test). **f**, Proposed model for the co-adjuvant effects of RA and IL-15 in the intestinal mucosa. Under inflammatory conditions, the expression of the proinflammatory cytokine IL-15 is upregulated in the lamina propria of the small intestine. Through the synergistic action of IL-15 and RA, dendritic cells acquire the ability to release inflammatory cytokines, particularly IL-12 and IL-23. These inflammatory mediators then act in concert with RA to prevent the induction of Foxp3⁺ T_{reg} cells and drive T_H1 and potentially T_H17 polarization. In turn, inflammatory T cells may provide help to B cells to produce specific IgG and IgA antibodies.

of inflammatory anti-gluten immunity develops without microbial adjuvant or systemic immunization in immunologically competent mice with a polyclonal T-cell-receptor repertoire. However, further studies are warranted to establish whether the anti-TG2 antibodies are gluten dependent. Especially relevant to coeliac disease is the

identification of IL-15 in the lamina propria as a causative factor driving the differentiation of anti-gluten CD4⁺ and CD8⁺ T_H1 cells in the intestinal mucosa, resulting in the break of tolerance to gluten. Our observations may also explain why oral tolerance is disrupted in patients with inflammatory bowel disease²⁸ who also have dysregulated

IL-15 expression in the gut²⁹. Lastly, our results indicate that inhibiting IL-15 signalling may constitute a therapeutic intervention to restore mucosal tolerance to luminal antigens.

METHODS SUMMARY

All knockout and transgenic mice used in these studies are on the C57BL/6 background. D^d-IL-15 transgenic mice⁹ expressing IL-15 under the minimal MHC class I D^d promoter were used to perform the studies. To study the response to dietary gluten, D^d-IL-15 transgenic mice were crossed to humanized HLA-DQ8 mice¹⁹ that had been backcrossed for twelve generations to MHC class II^{-/-} C57BL/6 mice. To assess T_{reg} conversion *in vivo*, RAG^{-/-} OT-II CD25⁻ CD4⁺ T cells were adoptively transferred into congenic OVA-fed C57BL/6 and D^d-IL-15 transgenic mice. For *in vivo* experiments mice were fed OVA, α -gliadin, or crude gliadin. T_{reg} differentiation assays were performed using CD4⁺eGFP⁺ T cells isolated from Foxp3^{eGFP} reporter mice and MLN or splenic dendritic cells from IL-15 transgenic or wild-type mice stimulated with IL-15, TGF- β and RA. To assess the role of IL-2R β , JNK2 or IL-12p40 on T_{reg}- and T_H1-cell differentiation, dendritic cells deficient in these different factors were purified from the corresponding knockout mice. To assess the role of RA, experiments were performed either in the presence of the RAR antagonist (LE135), a RAR α agonist, or with RAR α ^{-/-} dendritic cells³⁰. JNK phosphorylation was analysed by western blot following dendritic-cell stimulation with IL-15 and RA. Lamina propria cells were isolated from intestinal biopsies of control and active coeliac disease patients to determine levels of IL-12 and IL-15 expression by ELISA.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 March 2010; accepted 19 January 2011.

Published online 9 February 2011.

- Curotto de Lafaille, M. A. & Lafaille, J. J. Natural and adaptive Foxp3⁺ regulatory T cells: more of the same or a division of labor? *Immunity* **30**, 626–635 (2009).
- Faria, A. M. & Weiner, H. L. Oral tolerance. *Immunol. Rev.* **206**, 232–259 (2005).
- Jabri, B. & Sollid, L. M. Tissue-mediated control of immunopathology in coeliac disease. *Nature Rev. Immunol.* **9**, 858–870 (2009).
- Coombes, J. L. *et al.* A functionally specialized population of mucosal CD103⁺ DCs induces Foxp3⁺ regulatory T cells via a TGF- β and retinoic acid-dependent mechanism. *J. Exp. Med.* **204**, 1757–1764 (2007).
- Mora, J. R., Iwata, M. & von Andrian, U. H. Vitamin effects on the immune system: vitamins A and D take centre stage. *Nature Rev. Immunol.* **8**, 685–698 (2008).
- Mucida, D. *et al.* Reciprocal TH17 and regulatory T cell differentiation mediated by retinoic acid. *Science* **317**, 256–260 (2007).
- Mention, J. J. *et al.* Interleukin 15: a key to disrupted intraepithelial lymphocyte homeostasis and lymphomagenesis in coeliac disease. *Gastroenterology* **125**, 730–745 (2003).
- Tagaya, Y., Bamford, R. N., DeFilippis, A. P. & Waldmann, T. A. IL-15: a pleiotropic cytokine with diverse receptor/signaling pathways whose expression is controlled at multiple levels. *Immunity* **4**, 329–336 (1996).
- Fehniger, T. A. *et al.* Fatal leukemia in interleukin-15 transgenic mice follows early expansions in natural killer and memory phenotype CD8⁺ T cells. *J. Exp. Med.* **193**, 219–232 (2001).
- Caretto, D. *et al.* Cutting edge: the Th1 response inhibits the generation of peripheral regulatory T cells. *J. Immunol.* **184**, 30–34 (2010).
- Yokoyama, S. *et al.* Antibody-mediated blockade of IL-15 reverses the autoimmune intestinal damage in transgenic mice that overexpress IL-15 in enterocytes. *Proc. Natl Acad. Sci. USA* **106**, 15849–15854 (2009).
- Bettelli, E. *et al.* Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. *Nature* **441**, 235–238 (2006).
- Gao, Y., Camacho, L. H. & Mehta, K. Retinoic acid-induced CD38 antigen promotes leukemia cells attachment and interferon- γ /interleukin-1 β -dependent apoptosis of endothelial cells: implications in the etiology of retinoic acid syndrome. *Leuk. Res.* **31**, 455–463 (2007).
- Mohty, M. *et al.* All-trans retinoic acid skews monocyte differentiation into interleukin-12-secreting dendritic-like cells. *Br. J. Haematol.* **122**, 829–836 (2003).
- Rochette-Egly, C. & Germain, P. Dynamic and combinatorial control of gene expression by nuclear retinoic acid receptors (RARs). *Nucl. Recept. Signal.* **7**, e005 (2009).
- Mazzarella, G. *et al.* Gliadin activates HLA class I-restricted CD8⁺ T cells in coeliac disease intestinal mucosa and induces the enterocyte apoptosis. *Gastroenterology* **134**, 1017–1027 (2008).
- Nilsen, E. M. *et al.* Gluten specific, HLA-DQ restricted T cells from coeliac mucosa produce cytokines with Th1 or Th0 profile dominated by interferon gamma. *Gut* **37**, 766–776 (1995).
- Dieterich, W. *et al.* Identification of tissue transglutaminase as the autoantigen of coeliac disease. *Nature Med.* **3**, 797–801 (1997).
- Black, K. E., Murray, J. A. & David, C. S. HLA-DQ determines the response to exogenous wheat proteins: a model of gluten sensitivity in transgenic knockout mice. *J. Immunol.* **169**, 5595–5600 (2002).
- Marsh, M. N. Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology* **102**, 330–354 (1992).
- Meresse, B. *et al.* Coordinated induction by IL15 of a TCR-independent NKG2D signaling pathway converts CTL into lymphokine-activated killer cells in coeliac disease. *Immunity* **21**, 357–366 (2004).
- Monteleone, I. *et al.* Characterization of IL-17A-producing cells in coeliac disease mucosa. *J. Immunol.* **184**, 2211–2218 (2010).
- Bodd, M. *et al.* HLA-DQ2-restricted gluten-reactive T cells produce IL-21 but not IL-17 or IL-22. *Mucosal Immunol.* **3**, 594–601 (2010).
- Hunt, K. A. *et al.* Newly identified genetic risk variants for coeliac disease related to the immune response. *Nature Genet.* **40**, 395–402 (2008).
- Reddy, D., Siegel, C. A., Sands, B. E. & Kane, S. Possible association between isotretinoin and inflammatory bowel disease. *Am. J. Gastroenterol.* **101**, 1569–1573 (2006).
- Stephensen, C. B. & Livingston, K. A. Vitamin supplements and vaccines: maximize benefits, evaluate potential risks. *Am. J. Clin. Nutr.* **90**, 457–458 (2009).
- Holmgren, J. & Czernikinsky, C. Mucosal immunity and vaccines. *Nature Med.* **11** (suppl.), S45–S53 (2005).
- Kraus, T. A. *et al.* Failure to induce oral tolerance to a soluble protein in patients with inflammatory bowel disease. *Gastroenterology* **126**, 1771–1778 (2004).
- Liu, Z. *et al.* IL-15 is highly expressed in inflammatory bowel disease and regulates local T cell-dependent cytokine production. *J. Immunol.* **164**, 3608–3615 (2000).
- Hill, J. A. *et al.* Retinoic acid enhances Foxp3 induction indirectly by relieving inhibition from CD4⁺CD44^{hi} cells. *Immunity* **29**, 758–770 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank coeliac disease patients and their family members as well as the University of Chicago Celiac Disease Center for supporting our research. We thank B. Sally, L. M. Sollid and M. Musch for critical reading of the manuscript. We thank C. Ciszewski, B. Uzunparmak, and N. Grandison for their help with the collection and analysis of human biopsies. We thank M. Constantinides for technical assistance with mice breeding. We also thank the University of Chicago flow cytometry facility for technical assistance. D^d-IL-15 transgenic mice were a gift from M. Caligiuri. RAR α -deficient mice were provided by P. Chambon and C. Benoist. This work was supported by the Digestive Disease Research Core Center at the University of Chicago (DK42086), R01 DK67180 (for B.J.), R01DK71003 (for J.A.M.), and the Crohn's and Colitis Foundation (for V.A.).

Author Contributions R.W.D. and V.A. provided input into the conceptual development and execution of the studies, as well as preparation of the manuscript. F.T., H.F.-P., J.A.H. and W.W. provided technical assistance and input into data analyses. J.A.M. and E.V.M. helped with the analysis of the humanized HLA-DQ8 transgenic mice. D.D.K. provided preparations of α -gliadin used in the feeding experiments. T.A.W. provided TM β -1 antibody, and Y.B. helped with the realization of T-cell transfer experiments and provided us with RAR α -deficient bone-marrow. C.S., S.K. and S.G. followed patients with coeliac disease and provided intestinal biopsies for cytokines analysis. Y.B., J.A.M., D.D.K. and T.A.W. participated in discussion and review of the manuscript. B.J. conceived the idea, wrote the manuscript and supervised all investigations.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.J. (bjabri@bsd.uchicago.edu).

METHODS

Mice. C57BL/6 wild-type, IL-12p40^{-/-} and IL-2Rβ^{-/-} mice were purchased from Jackson Laboratory. RAG^{-/-} OT-II mice were bred and housed in our animal facility. Foxp3^{eGFP} reporter mice were previously described³¹. JNK2^{-/-} mice were provided by A. Lin. D^d-IL-15 transgenic mice⁹ expressing IL-15 under the minimal MHC class I promoter were donated by M. Caligiuri and crossed onto Ly5.1⁺ C57BL/6 mice or crossed onto HLA-DQ8 mice¹⁹, which were previously bred on an MHC class II^{-/-} C57BL/6 background. All mice were kept under specific pathogen-free conditions at the animal facility of the University of Chicago. All mice were maintained on standard rodent chow; however, DQ8 and DQ8-D^d-IL-15 transgenic mice were maintained on gluten-free chow obtained from Research Diets (AIN-76A). All experiments were performed in accordance with the Institutional Biosafety Committee and the Institutional Care and Use Committee.

T-cell and dendritic-cell purification. For CD4⁺ eGFP⁻ T-cell isolation, spleens and peripheral lymph nodes were mechanically disrupted through a 70-μm cell strainer. CD4⁺ cells were isolated by positive immunoselection using CD4-(L3T4) microbeads (Miltenyi Biotec). Purified CD4 T cells were sorted for GFP expression using BD FACSria (BD Bioscience). In some experiments, purified CD4⁺ CD44^{lo} GFP⁻ T cells were sorted using the FACSria (BD Bioscience).

For dendritic-cell isolation, MLN and spleen were digested with 400 units ml⁻¹ collagenase type IV (Sigma-Aldrich). Cells were filtered, resuspended in 22.5% Optiprep (Sigma-Aldrich), overlaid with Hank's Buffered Saline (HBS) and centrifuged at 670g for 30 min. Dendritic cells were then enriched from the interface by positive immunomagnetic selection using anti-CD11c-coated beads according to the manufacturer's recommendations (Miltenyi Biotec). Purification yielded up to 90% CD11c⁺ cells.

In vitro T_{reg}-cell-differentiation assay. For dendritic-cell-free cultures, 2 × 10⁵ CD4⁺ eGFP⁻ T cells were cultured for three days with 1 μg ml⁻¹ plate-bound anti-CD3ε (eBioscience) and 2 μg ml⁻¹ anti-CD28 (eBioscience). For cultures containing dendritic cells, 1 × 10⁵ CD4⁺ eGFP⁻ T cells and 4 × 10⁴ purified splenic or MLN dendritic cells were cultured for three days with 1 μg ml⁻¹ plate-bound anti-CD3ε (eBioscience). Specified recombinant cytokines (listed below) or neutralizing antibodies were added to the cultures.

Recombinant cytokines used were: TGF-β (2 μg ml⁻¹); IL-15 (20 ng ml⁻¹); IL-12p70 (250 pg ml⁻¹ to 2 ng ml⁻¹); IL-23 (250 pg ml⁻¹ to 2 ng ml⁻¹) (R&D). Neutralizing IL-12p40 antibody (1 μg ml⁻¹) (R&D). RA (10 nM) (Sigma-Aldrich); RAR antagonist LE135 (1 μM) (Tocris Bioscience).

T_H17 polarization. T_H17 differentiation was performed as previously described³². Briefly, purified CD4⁺ T cells were stimulated with plate bound anti-CD3ε (1 μg ml⁻¹) and anti-CD28 (2 μg ml⁻¹) in the presence of 2 ng ml⁻¹ TGF-β, 20 ng ml⁻¹ IL-6 in the presence or absence of 10 nM RA and 20 ng ml⁻¹ IL-15.

Bone-marrow-derived dendritic-cell culture. Culture of bone-marrow-derived dendritic cells was performed as previously described³³. 10⁵ bone-marrow-derived dendritic cells from RARα^{-/-} (ref. 34) or wild-type mice were cultured overnight with IL-15 (20 ng ml⁻¹) and RA (10 nM) or IL-15 (20 ng ml⁻¹) and RARα agonist AM580 (10 nM) (Tocris).

Antibodies and flow cytometry. The following conjugated antibodies were purchased from eBioscience (San Diego): CD4 (GK1.5), CD11c (N418), CD8β (eBioH35-17.2), TCRβ (H57-597), CD45.1 (A20), CD45.2 (104), IFN-γ (XMGI.2), IL-17 (eBio17B7), Foxp3 (FJK-16a) and isotype controls. The following antibodies were purchased from BD Biosciences: CD44 (IM7), H-2D^b (KH95), H-2D^d (34-2-12), H-2K^b (AF6-88.5), CD45 (30-F11), CD13 (R3-242) and isotype controls. Cells were permeabilized with the CytoFix/CytoPerm kit (BD Biosciences) for intra-cytoplasmic detection of IFN-γ and IL-17 cytokines. Foxp3 fixation/permeabilization kit was used for intranuclear detection of Foxp3 (eBioscience). Flow cytometry analysis was performed with a FACScanto (BD Biosciences).

Preparation of conditioned media and stimulation of T-cell cultures. 10⁵ splenic CD11c⁺ cells were stimulated with 10 nM RA, 1 μM LE135 and/or 20 ng ml⁻¹ rIL-15 and cultured for 24 h. One hundred and fifty microlitres of the supernatants were added to 2 × 10⁵ CD4⁺ Foxp3^{eGFP} T cells and differentiated on anti-CD3ε- and anti-CD28-coated plates for 72 h as described earlier.

IEL and lamina propria lymphocytes isolation. IELs³⁵ and lamina propria lymphocytes³⁶ were isolated as previously described using EDTA containing calcium-free media and collagenase VIII, respectively.

Cell signalling. Wild-type or IL-2Rβ^{-/-} dendritic cells were serum deprived overnight before stimulation with indicated doses of IL-15 or IL-15 and RA for 15 min. Preparation of cells and protein was performed as previously described³³. Before stimulation with IL-15, bone-marrow-derived dendritic cells were

pre-incubated for 30 min with various concentrations of MAP kinase inhibitors specific for MEK1/2 (PD98059), JNK (SP600125) and P38 (SB203580). In some experiments bone-marrow-derived dendritic cells were pre-incubated with cyclohexamide (10 μg ml⁻¹) (Sigma) or actinomycin D (10 μg ml⁻¹) (Sigma) for 15 min before stimulation with RA (0.1 ng ml⁻¹). Cells treated with DMSO were used as control for cells treated with MAP kinase inhibitors, cyclohexamide and actinomycin D.

T-cell transfers. CD4⁺ CD25⁻ T cells were purified from the spleen of RAG^{-/-} OT-II Ly5.1 or RAG^{-/-} OT-II Ly5.2 mice. 5 × 10⁵ cells were then transferred intravenously into naive Ly5.1 or Ly5.2 C57BL/6 and D^d-IL-15 transgenic mice. One day after transfer the mice received OVA dissolved in drinking water for five days or were administered 100 μg OVA every day for ten days by gavage as indicated in the figure legends. Mice were killed one day after the last feeding and intranuclear levels of Foxp3 were evaluated by flow cytometry in the transferred T cells.

Antigen feeding and immunization. Mice were fed by intragastric gavage using an 18-gauge round-tipped needle (Kent Scientific) with the following proteins: 100 μg ovalbumin (OVA) (Sigma-Aldrich), 100 μg recombinant α-gliadin³⁷ dissolved in water, or 20 mg crude gliadin (Sigma-Aldrich). In some experiments, mice were also fed with LE135 (1 μM) or RA (1 μM) resuspended in corn oil. Feeding occurred every other day for ten days. In some experiments, D^d-IL-15 transgenic mice were injected i.p. with 100 μg of anti-IL-12p40 (R&D systems, clone C17.8, rat IgG2a) at the time of feedings, or with 20 μg of anti-mouse IL-15 antibody (M96, Amgen, mouse IgG2a) once a week before and during feeding, or 200 μg of a purified anti-mouse CD122 antibody (clone TM-β1, rat IgG2b) twice a week before and during feeding.

One day following the last feeding, lamina propria and MLN cells were isolated and re-stimulated for 24 h or 48 h, respectively with 50 μg ml⁻¹ OVA or gliadin as noted in the figure legend. Supernatants were analysed for cytokines by ELISA.

Detection of cytokines by ELISA. Cell supernatants were evaluated for IL-23 (R&D), IL-12p70 (BD Biosciences), IL-6, (BD Biosciences), IFN-γ (BD Biosciences) and IL-17 (R&D). Tissue pellets from epithelium, lamina propria and MLN were quantified for IL-15 (eBioscience).

Anti-TG2 and anti-gluten ELISA. Serum was harvested 15 days after mice received last gluten feeding. ELISA assays were performed as previously described³⁸.

Patients and controls. Seven patients (age: 3–46 years) with active coeliac disease were investigated. Diagnosis of coeliac disease was based on the detection of anti-transglutaminase antibodies, the expression of HLA DQ2 or DQ8, villous atrophy, and clinical and histological response to a gluten-free diet. Fourteen individuals (age: 1–55 years) undergoing endoscopies and biopsies for functional intestinal disorders of non-coeliac origin were studied as controls. Lamina propria cells were isolated from biopsies or surgical specimens as previously described²³. All subjects gave written informed consent, and research was approved by the institutional review boards. Lamina propria cells were centrifuged and resuspended in 500 ml PBS. Protein concentration was determined and total protein was analysed for levels of IL-15 (eBioscience) and IL-12p70 (BD Biosciences).

Histology. Haematoxylin & eosin staining was performed on 5 μM 10% formalin-fixed paraffin-embedded intestinal sections. Slides were analysed under a Leica DM 2500 microscope with a HC PLAN APO ×20/0.7 NA and a HCX PL APO ×100/1.40–0.70 objectives.

- Fontenot, J. D. *et al.* Regulatory T cell lineage specification by the forkhead transcription factor Foxp3. *Immunity* **22**, 329–341 (2005).
- Veldhoen, M. *et al.* TGFβ in the context of an inflammatory cytokine milieu supports *de novo* differentiation of IL-17-producing T cells. *Immunity* **24**, 179–189 (2006).
- Depaolo, R. W. *et al.* Toll-like receptor 6 drives differentiation of tolerogenic dendritic cells and contributes to LcrV-mediated plague pathogenesis. *Cell Host Microbe* **4**, 350–361 (2008).
- Hill, J. A. *et al.* Retinoic acid enhances Foxp3 induction indirectly by relieving inhibition from CD4⁺ CD44^{hi} cells. *Immunity* **29**, 758–770 (2008).
- Park, S. H. *et al.* Selection and expansion of CD8α/α¹ T cell receptor α/β¹ intestinal intraepithelial lymphocytes in the absence of both classical major histocompatibility complex class I and nonclassical CD1 molecules. *J. Exp. Med.* **190**, 885–890 (1999).
- Lefrançois, L. & Lycke, N. Isolation of mouse small intestinal intraepithelial lymphocytes, Peyer's patch, and lamina propria cells. *Curr. Protoc. Immunol.* doi:10.1002/0471142735.im0319s17 (2001).
- Bernardin, J. E., Kasarda, D. D. & Mechem, D. K. Preparation and characterization of α-gliadin. *J. Biol. Chem.* **242**, 445–450 (1967).
- Marietta, E. *et al.* A new model for dermatitis herpetiformis that uses HLA-DQ8 transgenic NOD mice. *J. Clin. Invest.* **114**, 1090–1097 (2004).

Modelling the long QT syndrome with induced pluripotent stem cells

Ilanit Itzhaki^{1*}, Leonid Maizels^{1*}, Irit Huber^{1*}, Limor Zwi-Dantsis¹, Oren Caspi¹, Aaron Winterstern¹, Oren Feldman¹, Amira Gepstein¹, Gil Arbel¹, Haim Hammerman², Monther Boulos² & Lior Gepstein^{1,2}

The ability to generate patient-specific human induced pluripotent stem cells (iPSCs)^{1–3} offers a new paradigm for modelling human disease and for individualizing drug testing⁴. Congenital long QT syndrome (LQTS) is a familial arrhythmogenic syndrome characterized by abnormal ion channel function and sudden cardiac death^{5–7}. Here we report the development of a patient/disease-specific human iPSC line from a patient with type-2 LQTS (which is due to the A614V missense mutation in the *KCNH2* gene). The generated iPSCs were coaxed to differentiate into the cardiac lineage. Detailed whole-cell patch-clamp and extracellular multielectrode recordings revealed significant prolongation of the action-potential duration in LQTS human iPSC-derived cardiomyocytes (the characteristic LQTS phenotype) when compared to healthy control cells. Voltage-clamp studies confirmed that this action-potential-duration prolongation stems from a significant reduction of the cardiac potassium current I_{Kr} . Importantly, LQTS-derived cells also showed marked arrhythmogenicity, characterized by early-after depolarizations and triggered arrhythmias. We then used the LQTS human iPSC-derived cardiac-tissue model to evaluate the potency of existing and novel pharmacological agents that may either aggravate (potassium-channel blockers) or ameliorate (calcium-channel blockers, K_{ATP} -channel openers and late sodium-channel blockers) the disease phenotype. Our study illustrates the ability of human iPSC technology to model the abnormal functional phenotype of an inherited cardiac disorder and to identify potential new therapeutic agents. As such, it represents a promising paradigm to study disease mechanisms, optimize patient care (personalized medicine), and aid in the development of new therapies.

Congenital LQTS^{5–7}, now classified into 12 subtypes, is a familial arrhythmogenic syndrome characterized by delayed repolarization, a prolonged QT interval in the electrocardiogram and a life-threatening polymorphic ventricular tachycardia known as torsade de pointes (TdP). Although heterologous expression systems⁶ and animal models have provided important insights into LQTS pathogenesis, the lack of *in vitro* sources for human cardiomyocytes and the inability to model patient-specific disease variations has significantly hampered the study of this disease. Here we propose that cardiomyocytes derived from LQTS patient-specific human iPSC lines may recapitulate the disease phenotype *in vitro*, providing important mechanistic insights and offering a unique platform to evaluate patient-specific disease aggravators and therapies (Supplementary Fig. 1).

Dermal fibroblasts were obtained from a 28-year-old woman with a diagnosis of familial type-2 LQTS (Fig. 1a) due to a missense mutation in exon 9 of the *KCNH2* gene (A614V mutation caused by a C→T nucleotide substitution)⁸. This mutation was previously shown in heterologous expression systems to affect the pore-forming region of the *KCNH2* (also known as *HERG*)-encoded potassium channel (Supplementary Fig. 1); leading to a significant reduction of the rapid component of the delayed-rectifier potassium current (I_{Kr}) responsible for type-2 LQTS⁹.

The cultured fibroblasts were reprogrammed to generate LQTS patient-specific human iPSC clones (Fig. 1b) after transduction with retroviral vectors encoding for SOX2, KLF4 and OCT4. Several clones were generated, three of which were continuously propagated and used for cardiomyocyte differentiation and characterization. Similar control human iPSCs were created from fibroblasts from a healthy individual. Importantly, the A614V heterozygous mutation was identified in LQTS human iPSCs (Fig. 1a) but not in control iPSCs.

All iPSC clones generated showed characteristic human embryonic stem (ES) cell morphology (Fig. 1b and Supplementary Fig. 2), expressed the pluripotency markers NANOG, SSEA4, OCT4 and TRA-1-60 (Fig. 1b and Supplementary Fig. 2), had alkaline phosphatase activity (Fig. 1b and Supplementary Fig. 3A), and maintained a normal karyotype (Fig. 1c and Supplementary Fig. 3B). All clones showed silencing of the three retroviral transgenes (Supplementary Fig. 4), reactivation of endogenous pluripotency genes (*OCT4*, *SOX2*, *NANOG*, *FOXD3* and *ZFP42* (also known as *REX1*); Supplementary Fig. 5), and demethylation of the *NANOG* promoter regions (Fig. 1d); all indicating successful reprogramming. Pluripotency of the human iPSC clones was confirmed by the presence of cell derivatives of all three germ layers in *in-vitro*-differentiating embryoid bodies (Fig. 1e and Supplementary Fig. 6) and by teratoma formation after injection of undifferentiated human iPSCs into immunocompromised NOD/SCID mice (Fig. 1f and Supplementary Fig. 7).

Next we used the embryoid-body-differentiation system to coax human iPSC differentiation into the cardiac lineage. Similarly to human ES cells^{10,11} and healthy human iPSCs^{12,13}, this resulted in the appearance of spontaneously beating areas within the differentiating embryoid bodies (Supplementary Movies 1 and 2). Both the control and LQTS human iPSC lines showed comparable cardiomyocyte differentiation capacities (Supplementary Fig. 8). Immunocytochemical studies showed positive staining of the control (Supplementary Fig. 9A) and LQTS human iPSC-derived cardiomyocytes (iPSC-CMs) for the sarcomeric proteins troponin I (TnI) and α -actinin and for the gap-junction protein connexin 43 (Fig. 2a). Similarly, gene expression analysis confirmed the expression of cardiac-specific transcription factors (*NKX2-5*) and structural genes (*MLC2V* (also known as *MYL2*), *MYH6* and *MYH7*) by both cell types, including of the relevant *KCNH2* gene (Fig. 2b). Importantly, both control and LQTS iPSC-CMs showed cardiac-specific functional properties. This was manifested by the development of action potentials at the cellular level (Fig. 3a), a functional syncytium at the multicellular level (Fig. 2c and Supplementary Fig. 9B) and appropriate chronotropic responses to the β -adrenergic agonist isoproterenol (Fig. 2d and Supplementary Fig. 9C).

Similar to previous reports in human ES cells¹⁰ and human iPSCs¹², three types of action-potential morphologies were recorded from control and LQTS iPSC-CMs (Fig. 3a and Supplementary Table 1). The most dominant form was the 'ventricular-like' morphology (characterized by a significant plateau phase), with the 'atrial-like' (showing

¹Sohnis Family Research Laboratory for Cardiac Electrophysiology and Regenerative Medicine, the Bruce Rappaport Faculty of Medicine, Technion — Israel Institute of Technology, POB 9649, Haifa 31096, Israel. ²Department of Cardiology, Rambam Medical Center, the Bruce Rappaport Faculty of Medicine, Technion — Israel Institute of Technology, POB 9649, Haifa 31096, Israel.

*These authors contributed equally to this work.

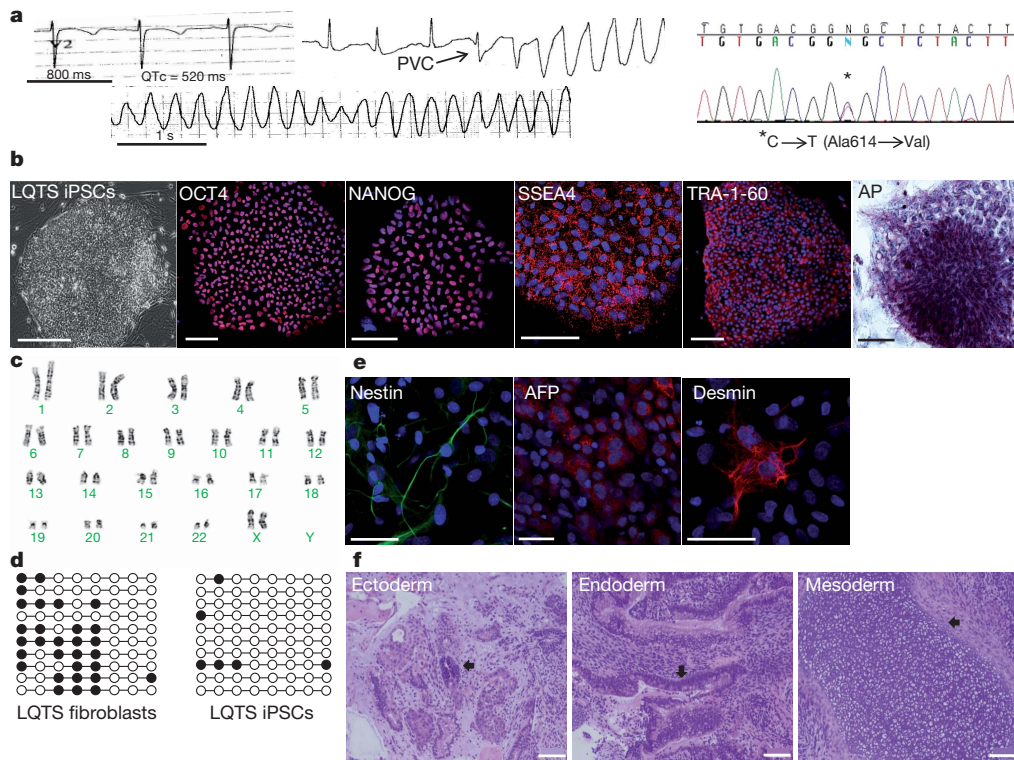


Figure 1 | Establishment and characterization of type-2 LQTS human iPSCs (clone 1). **a**, Left panel shows ECGs from the type-2 LQTS patient during sinus rhythm (QT interval corrected for heart rate (QTc) = 520 ms) and during initiation and sustainment of TdP. PVC, premature ventricular contraction. Right panel shows sequencing of the *KCNH2* gene identifying the A614V heterozygous missense mutation in LQTS human iPSCs. **b**, Reprogramming of the patient's fibroblasts into LQTS human iPSCs. Shown from left to right are a typical human iPSC colony, positive immunostaining for pluripotency markers (OCT4, NANOG, SSEA4 and TRA-1-60), and positive staining for alkaline

phosphatase (AP) activity. Scale bars, 100 μm. **c**, Karyotype analysis of LQTS human iPSCs. **d**, Bisulphite sequencing analysis of the NANOG promoter in the patient's fibroblasts and LQTS human iPSCs. **e**, Immunostaining of *in vitro* differentiating embryoid bodies for nestin (ectoderm), α-fetoprotein (AFP, endoderm) and desmin (mesoderm). Scale bars, 50 μm. **f**, Teratoma formation following injection of undifferentiated LQTS human iPSCs in NOD/SCID mice. Note the formation of pigmented epithelium (ectoderm), gastrointestinal epithelium (endoderm) and hyaline cartilage (mesoderm), as identified by the arrows. Scale bars, 200 μm.

triangular-shaped action potentials) and 'nodal-like' (showing relatively depolarized resting membrane potential) types being less prevalent. More importantly, these intracellular recordings revealed marked action-potential duration (APD) prolongation in LQTS iPSC-CMs (the characteristic LQTS electrophysiological signature⁵⁻⁷) when compared to control cells (Fig. 3a, b and Supplementary Table 1). APD prolongation and the associated reduced repolarization velocity were noted in both ventricular-like and atrial-like LQTS human iPSC-CMs but not in the nodal-like cells.

Because APD is rate-dependent, next we electrically stimulated the human iPSC-CMs and noted that the differences in APD values between LQTS and control cells persisted also at fixed rates (0.5 and 1 Hz; Fig. 3b and Supplementary Fig. 10). In addition, comparable prolonged APD values were noted in ventricular myocytes derived from three different independent LQTS iPSC clones, which were all significantly longer than those obtained from two different control human iPSC lines (Supplementary Fig. 11A). Lastly, the LQTS phenotype could be recapitulated in control iPSC-CMs by pharmacologically blocking the I_{K_r} current with the specific blocker E-4031 (Supplementary Fig. 12).

Next we performed single-cell voltage-clamp studies and identified the presence of an E4031-sensitive current (I_{K_r}) in control human iPSC-CMs (Fig. 3c), which resembled previous I_{K_r} recordings from explanted human ventricular myocytes¹⁴. Similar recordings in the LQTS iPSC-CMs revealed significant reductions of this current (Fig. 3c). Thus, peak amplitudes of the I_{K_r} activation currents in LQTS cardiomyocytes ($n = 6$), measured at physiologically relevant depolarization steps (0, 20 mV), were found to be significantly smaller (by 72 and 60%, respectively) than those recorded in control cells ($P < 0.05$; Fig. 3d, left).

Similarly, peak amplitudes of the I_{K_r} tail currents measured at 20 and 60 mV were also significantly smaller (by 64 and 68%, respectively, $P < 0.05$; Fig. 3d, right) than in control cells. These results are consistent with previous recordings in heterologous expression systems, where co-injection of wild-type along with mutant A614V cRNA suppressed the HERG current in a dominant-negative manner⁹.

To evaluate the electrophysiological properties also at the multicellular level, we studied the human iPSC-derived cardiac tissue with a microelectrode array mapping technique¹⁵. The recorded extracellular electrograms were analysed to measure the field-potential duration (FPD; Fig. 2e). This measurement, which was normalized to account for variations in beating frequency¹⁶ (corrected FPD, cFPD), is analogous to the QT interval in the electrocardiogram (ECG) and was previously shown to correlate with APD and to predict drug effects on repolarization^{15,17}. Similar to patch-clamp studies, cFPD (or rate-matched FPD) measurements were significantly longer in cardiac tissues derived from all three LQTS human iPSC clones when compared to healthy control specimens (Fig. 2e, f, Supplementary Fig. 11B and Supplementary Movie 1).

Importantly, LQTS iPSC-CMs showed marked arrhythmogenicity. This was manifested by the development of early-after depolarizations (EADs) in both the atrial-like and ventricular-like LQTS iPSC-CMs (Fig. 3e). EADs are spontaneous membrane depolarizations regarded as the harbinger of ventricular arrhythmias in LQTS⁷, and are thought to result from late membrane-inward currents, such as the L-type Ca^{2+} 'window' current¹⁸. When reaching threshold, EADs may give rise to triggered action potentials, manifested in LQTS iPSC-CMs as isolated premature beats or even multiple sequential triggered action potentials (Fig. 3f, g). The presence of EADs was observed in 38 of the 58 LQTS iPSC-CMs studied (66%). These EADs gave rise to a single premature

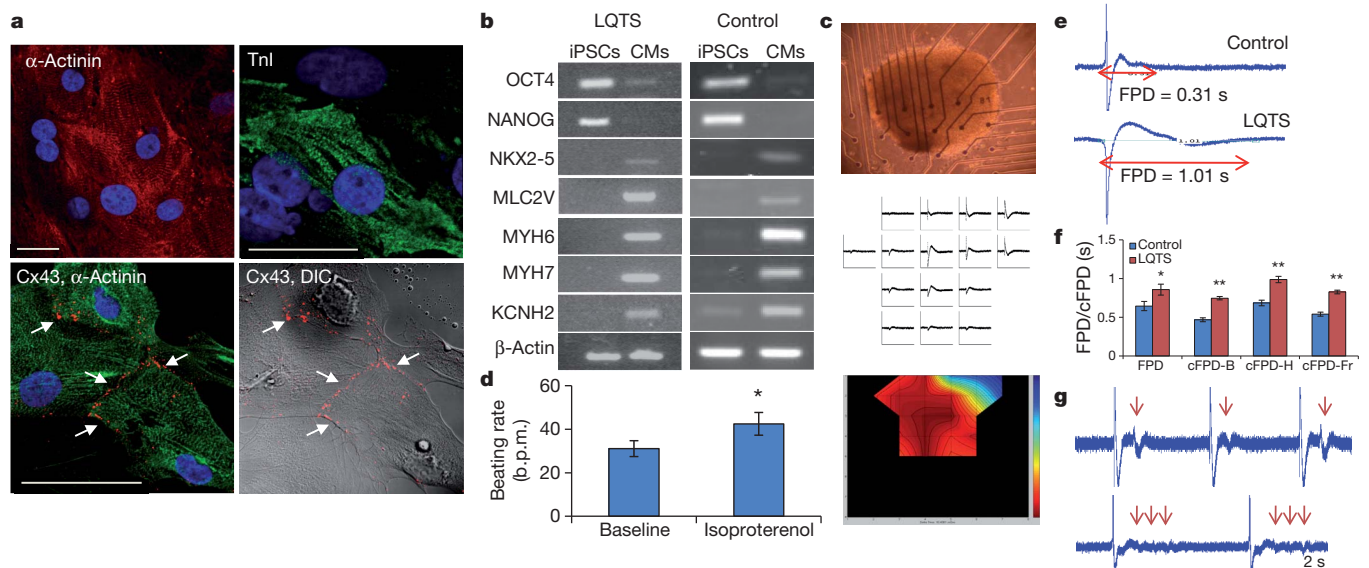


Figure 2 | Phenotypic characterization of LQTS human iPSC-derived cardiac-tissue. **a**, Immunostaining of LQTS iPSC-CMs for α -actinin and TnI (top). Gap junctions (arrows) were identified by the punctuated connexin 43 (Cx43) immunosignal (red) at the intercellular junctions (differential interference contrast microscopy (DIC) image, bottom right) between the cardiomyocytes (identified by α -actinin staining in green, bottom left). Scale bars, 30 μ m. **b**, Semi-quantitative polymerase chain reaction with reverse transcription (RT-PCR) demonstrating expression of cardiac-specific transcription factors (NKX2-5), sarcomeric proteins (MYH6, MYH7, MLC2V) and ion channels (KCNH2) by LQTS and control human iPSC-CMs in comparison to undifferentiated human iPSCs expressing OCT4 and NANOG. **c**, Microelectrode array extracellular recordings (middle) from LQTS human

iPSC cardiac tissue (top). The resulting colour-coded activation map (right) shows electrical propagation from the pacemaker area (red) to latest activation sites (blue). **d**, Positive chronotropic responses of LQTS hiPSCs-CMs ($n = 16$) to isoproterenol (10 μ M). * $P < 0.01$ when compared to baseline. b.p.m., beats per minute. **e**, Extracellular recordings from control (top) and LQTS (bottom) human iPSC-derived cardiac tissues and their respective FPD measurements. **f**, Summary of rate-matched FPD ($n = 12$) and cFPD values (using Bazett's (B), Hodges' (H) or Fridericia's¹⁶ (Fr) corrections) in control ($n = 23$) and LQTS human iPSC ($n = 45$) cardiac tissues. Error bars show s.e.m. * $P < 0.05$ and ** $P < 0.01$. **g**, Arrhythmogenicity in LQTS cardiac tissue identified as single (top) or multiple (bottom, arrows) premature beats.

beat in 36% of the cells and in 28% also to sustained triggered activity. In contrast, recordings from control iPSC-CMs ($n = 31$) failed to show any EADs or any other arrhythmogenic activity.

The arrhythmogenic activity of the LQTS iPSC-CMs was also identified at the multicellular level and ranged from isolated premature beats (Fig. 2g, top, and Supplementary Movie 2) to more sustained activity (Fig. 2g, bottom). Consequently, ectopic activity in the microelectrode array recordings was noted in 38% (26 of 84) of the LQTS human iPSC-derived cardiac tissues studied but only in 6% (3 of 47) of control samples. This ectopic activity was consistent with a triggered activity mechanism (although re-entry and automaticity could not be ruled out completely).

Next we used the LQTS human iPSC model to evaluate drugs that may either ameliorate or aggravate the disease phenotype¹⁹. Because malignant arrhythmias in LQTS patients are often precipitated by drugs that block I_{Kr} , we initially evaluated the specific I_{Kr} blocker E-4031 (500 nM), and noted significant APD prolongation (by $47 \pm 2\%$, $P < 0.01$, $n = 7$) in the LQTS iPSC-CMs (Supplementary Fig. 12B). This APD prolongation was coupled with increased arrhythmogenesis in 6 of the 7 cells studied, manifested either by the development of new EADs or by an increase in their number and complexity (Supplementary Fig. 12B). Similarly, I_{Kr} blockade by E-4031 or by cisapride (a gastric prokinetic agent previously removed from the market because of its indirect I_{Kr} blocking activity, which led to increased pro-arrhythmic mortality²⁰) significantly prolonged cFPD and increased arrhythmogenicity also at the multicellular level (Supplementary Fig. 12C, D). These results, together with the I_{Kr} current analysis, provide a measure of the magnitude of the dominant-negative effect induced by the A614V mutation. They also offer clues to the potential susceptibility of such LQTS patients to TdP in response to agents that may further inhibit the remaining I_{Kr} current. More broadly it highlights the potential role of

the human iPSC model for safety pharmacology²⁰, both for drug development and for patient-specific safety screening.

Next we used the LQTS human iPSC-CM model to study agents that may possess novel therapeutic benefits for LQTS, through either APD shortening or by direct EAD suppression¹⁹. The effect of these agents was also evaluated in healthy control cells (Supplementary Fig. 13). Because Ca^{2+} influx through L-type Ca^{2+} channels contributes importantly to APD and also has a role in EAD formation¹⁸, we proposed that inhibition of this current may be anti-arrhythmic¹⁹. Application of a Ca^{2+} -channel blocker (nifedipine, 1 μ M) led to significant APD abbreviation in LQTS cardiomyocytes (Fig. 4a, b, top), with mean APD₉₀ (the time interval required to reach 90% of repolarization) shortened by $57 \pm 7\%$ ($P < 0.01$). More importantly, nifedipine application resulted in complete elimination of all EADs and triggered beats (Fig. 4c; $n = 11$). Similarly, nifedipine (1 μ M) significantly shortened cFPD (Fig. 4a; $P < 0.01$, $n = 5$) and abolished all arrhythmic activity in the LQTS cardiac-tissue model (Fig. 4d). Long-term application of this agent, however, was associated with cessation of beating in some embryoid bodies.

An alternative strategy to favourably alter the balance between inward and outward currents during repolarization in the LQTS may be to augment outward potassium currents. To this end we evaluated the effects of pinacidil (1 μ M), a K_{ATP} -channel opener, and noted significant shortening of both APD₉₀ (Fig. 4a, b; $P < 0.05$, $n = 4$) and cFPD (Fig. 4a; $P < 0.05$, $n = 9$) in LQTS iPSC-CMs. Importantly, pinacidil application completely abolished all EADs and triggered arrhythmias in all LQTS iPSC-CMs studied (Fig. 4e, f; $n = 7$). A potential shortcoming of the aforementioned therapies may be excessive APD shortening leading to the potentially pro-arrhythmic 'short QT syndrome'¹⁹. In this respect, the human iPSC model may prove useful, as it may allow the identification of compounds and optimization of their dosage to result in APD normalization without pathological shortening.

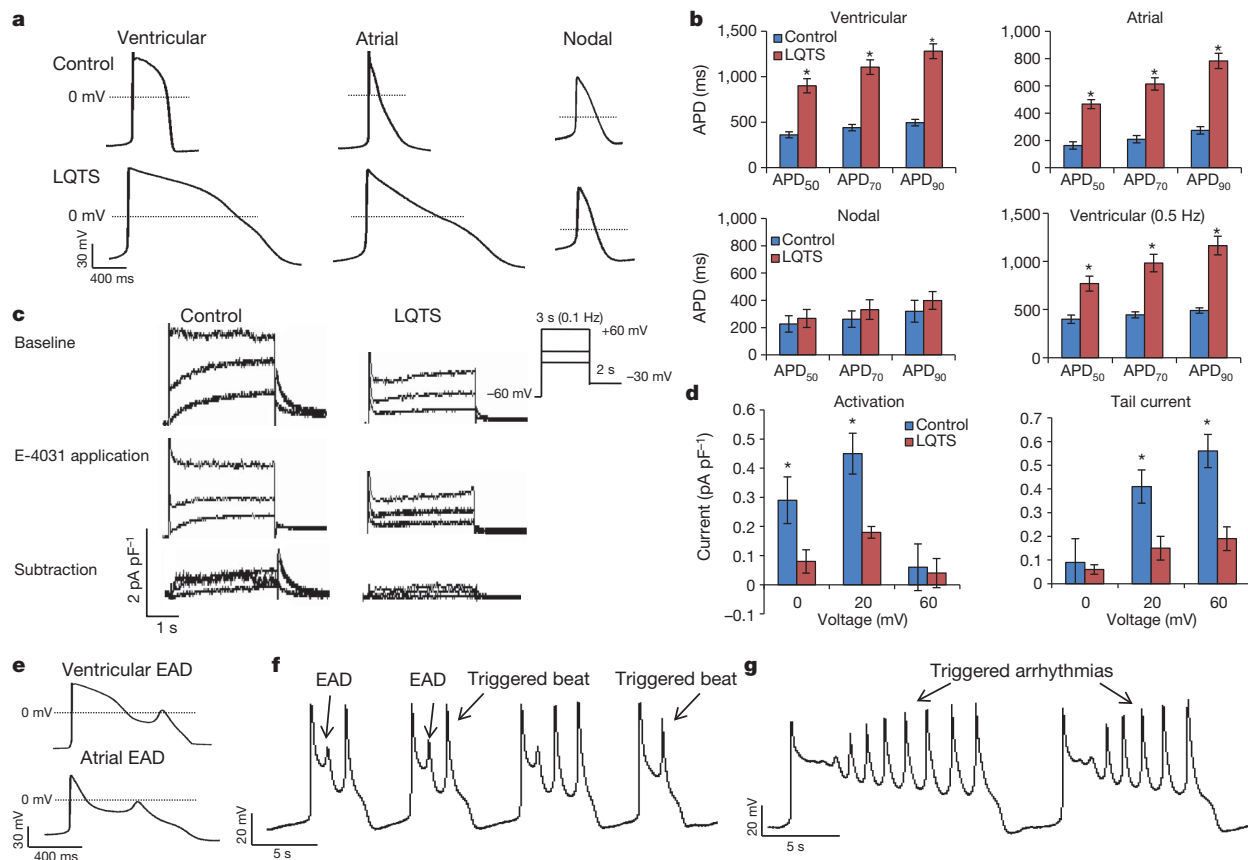


Figure 3 | Whole-cell patch-clamp recording from human iPSC-CMs. **a**, Action-potential recordings from control and LQTS human iPSC-CMs showing ventricular-like, atrial-like and nodal-like morphologies. Notice the marked APD prolongation in both ventricular-like and atrial-like LQTS iPSC-CMs. **b**, APD₉₀, APD₇₀, and APD₅₀ values in LQTS (red) and healthy (blue) human iPSC-CMs. **P* < 0.01 when compared to control cells. Error bars show s.e.m. **c**, Voltage-clamp recordings of the *I*_{Kr} current, measured from control (left) and LQTS (right) human iPSC-derived ventricular cardiomyocytes. Top, baseline recordings. Middle, recording following administration of E-4031

Lastly, we also evaluated the potential anti-arrhythmic efficacy of the late Na⁺-channel blocker, ranolazine²¹. Although, this agent may theoretically be more applicable for type-3 LQTS (resulting from gain-of-function Na⁺-channel mutations), a recent study showed its beneficial effects also in a canine model of acquired type-2 LQTS²². Interestingly, ranolazine application (15–50 μM) did not significantly alter APD₉₀ or cFPD in LQTS cardiomyocytes (Fig. 4a and Supplementary Fig. 14; *n* = 8), probably because of its nonspecific blocking effect on different ion channels. Nevertheless, ranolazine induced a pronounced anti-arrhythmic effect at both the cellular and multicellular levels (Fig. 4g, h and Supplementary Fig. 14).

Taken together, our data demonstrate the ability of human iPSC technology to model the abnormal functional phenotype of an inherited cardiac disorder. Using this approach, the disease phenotype presented by type-2 LQTS patients at the bedside (long QT interval and the development of TdP) could be recognized as a measurable anomalous electrophysiological signature in bench studies using patient-specific human iPSC-CMs. Our study adds to a recent report demonstrating the ability to model type-1 LQTS using human iPSCs²³. Importantly, the current study provides mechanistic insights into the pathogenesis of arrhythmias in this syndrome (by demonstrating the development of spontaneous EADs leading to triggered activity at both the cellular and multicellular levels). Moreover, it also highlights the potential of human iPSCs in the emerging field of personalized medicine by demonstrating the ability to screen the effects of potential disease aggravators and novel

(1 μM). Bottom, E-4031-sensitive current (*I*_{Kr}) defined by digital subtraction of the two currents. **d**, Summary of the *I*_{Kr} activation and tail-current amplitudes measured following test depolarization pulses of 0, 20 and 60 mV. Note the significant *I*_{Kr} attenuation in LQTS cells (red, *n* = 6) when compared to controls (blue, *n* = 6). Error bars show s.e.m. **e**, Development of EADs in LQTS ventricular-like and atrial-like cells. **f**, Development of triggered activity in LQTS cardiomyocytes, manifested as EADs or even single and multiple (g) triggered beats.

customized treatment options (which in LQTS include Ca²⁺-channel blockers, K_{ATP}-channel openers, and late Na⁺-channel blockers).

The results of the current study, focusing on congenital type-2 LQTS, may have broader implications, as the specific current affected (*I*_{Kr}) is thought to have an important role in the more commonly acquired LQTS forms (resulting from heart failure, cardiac hypertrophy, or drug therapy). Moreover, the concepts described here may also be extended to model several other human genetic disorders, enabling translational research into disease mechanisms and therapies.

METHODS SUMMARY

Human iPSC generation and cardiomyocyte differentiation. Dermal fibroblasts were reprogrammed to generate human iPSC clones by transduction with retroviral vectors encoding SOX2, KLF4 and OCT4, followed by valproic acid treatment²⁴. The undifferentiated human iPSC colonies were propagated on a MEF feeder layer. To induce differentiation, human iPSCs were dispersed into small cell clumps and cultivated in suspension for 10 days as embryoid bodies and then plated^{11,13}.

Immunostainings. Cells were fixed with 4% paraformaldehyde, permeabilized with 1% Triton, blocked with 5% horse serum, and incubated with primary antibodies (as listed in Methods). The preparations were incubated with secondary antibodies and counterstained with 4',6-diamidino-2-phenylindole (DAPI) for visualization.

Whole-cell patch-clamp recordings. Dispersed human iPSC-CMs, obtained by enzymatic dissociation (collagenase, 1 mg ml⁻¹) of contracting embryoid bodies, were attached to fibronectin-coated glass coverslips. Action potentials were recorded at 32 °C in the current-clamp mode. A custom-written Matlab program was used to analyse action-potential parameters. Voltage-clamp studies were performed to characterize the *I*_{Kr} current, defined as the E-4031 (1 μM) sensitive current. These

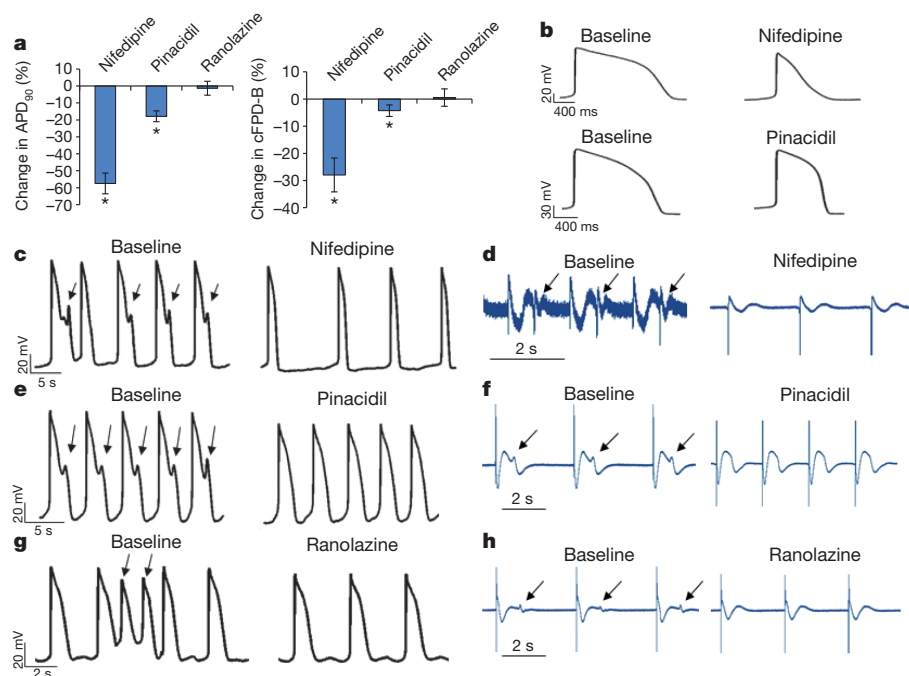


Figure 4 | Drug screening using LQTS human iPSC-CMs. **a**, Changes (percentage) in APD₉₀ and cFPD with Bazett's correction values following treatment with Ca²⁺-channel blocker nifedipine, K_{ATP}-channel opener pinacidil and late Na⁺-channel blocker ranolazine. *P < 0.05 when mean absolute values are compared to baseline. Error bars show s.e.m. **b**, APD shortening induced by nifedipine and pinacidil in LQTS iPSC-CMs. **c–h**, Anti-arrhythmic activity at the cellular (**c**, **e**, **g**) and multicellular (**d**, **f**, **h**) levels induced by nifedipine (**c**, **d**), pinacidil (**e**, **f**) and ranolazine (**g**, **h**). Arrows show the ectopic activity at baseline.

currents were elicited by 3-s depolarizing steps from a holding potential of -60 mV to potentials ranging from -40 to $+60$ mV in 10 mV increments. This was followed by a 2-s repolarization phase to -30 mV to elicit the tail current.

Microelectrode array recordings. Extracellular electrograms, recorded by a high-resolution microelectrode array recording system (multichannel systems)¹⁵, were used to determine FPD values in LQTS and control human iPSC-derived cardiac tissues. FPD was defined as the time interval between the initial field-potential deflection and its return to baseline. FPD measurements were either compared in rate-matched specimens or normalized (cFPD) for rate using Bazett's, Hodges' or Fridericia's¹⁶ correction formulas.

Statistical analysis. Results are reported as mean \pm s.e.m. Comparisons between LQTS human iPSC-CMs and healthy human iPSC-CMs groups were performed using the unpaired Student's *t*-test. Drug effects were analysed by paired *t*-test. One-way ANOVA followed by Tukey's post-hoc tests were used when comparing multiple groups. *P* < 0.05 was considered statistically significant.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 June; accepted 14 December 2010.

Published online 16 January 2011.

- Park, I. H. *et al.* Disease-specific induced pluripotent stem cells. *Cell* **134**, 877–886 (2008).
- Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
- Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
- Kiskinis, E. & Eggan, K. Progress toward the clinical application of patient-specific pluripotent stem cells. *J. Clin. Invest.* **120**, 51–59 (2010).
- Goldenberg, I. & Moss, A. J. Long QT syndrome. *J. Am. Coll. Cardiol.* **51**, 2291–2300 (2008).
- Sanguinetti, M. C. & Tristani-Firouzi, M. hERG potassium channels and cardiac arrhythmia. *Nature* **440**, 463–469 (2006).
- Marbán, E. Cardiac channelopathies. *Nature* **415**, 213–218 (2002).
- Tenenbaum, M. *et al.* Identification of the gene causing long QT syndrome in an Israeli family. *Isr. Med. Assoc. J.* **10**, 809–811 (2008).
- Nakajima, T. *et al.* Novel mechanism of HERG current suppression in LQT2: shift in voltage dependence of HERG inactivation. *Circ. Res.* **83**, 415–422 (1998).
- He, J. Q., Ma, Y., Lee, Y., Thomson, J. A. & Kamp, T. J. Human embryonic stem cells develop into multiple types of cardiac myocytes: action potential characterization. *Circ. Res.* **93**, 32–39 (2003).
- Kehat, I. *et al.* Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. *J. Clin. Invest.* **108**, 407–414 (2001).
- Zhang, J. *et al.* Functional cardiomyocytes derived from human induced pluripotent stem cells. *Circ. Res.* **104**, e30–e41 (2009).

- Zwi, L. *et al.* Cardiomyocyte differentiation of human induced pluripotent stem cells. *Circulation* **120**, 1513–1523 (2009).
- Li, G. R., Feng, J., Yue, L., Carrier, M. & Nattel, S. Evidence for two components of delayed rectifier K⁺ current in human ventricular myocytes. *Circ. Res.* **78**, 689–696 (1996).
- Caspi, O. *et al.* In vitro electrophysiological drug testing using human embryonic stem cell derived cardiomyocytes. *Stem Cells Dev.* **18**, 161–172 (2009).
- Luo, S., Michler, K., Johnston, P. & Macfarlane, P. W. A comparison of commonly used QT correction formulae: the effect of heart rate on the QTc of normal ECGs. *J. Electrocardiol.* **37** (suppl. 1), 81–90 (2004).
- Halbach, M., Egert, U., Hescheler, J. & Banach, K. Estimation of action potential changes from field potential recordings in multicellular mouse cardiac myocyte cultures. *Cell. Physiol. Biochem.* **13**, 271–284 (2003).
- January, C. T. & Riddle, J. M. Early afterdepolarizations: mechanism of induction and block. A role for L-type Ca²⁺ current. *Circ. Res.* **64**, 977–990 (1989).
- Patel, C. & Antzelevitch, C. Pharmacological approach to the treatment of long and short QT syndromes. *Pharmacol. Ther.* **118**, 138–151 (2008).
- Fermini, B. & Fossa, A. A. The impact of drug-induced QT interval prolongation on drug discovery and development. *Nature Rev. Drug Discov.* **2**, 439–447 (2003).
- Antzelevitch, C. *et al.* Electrophysiological effects of ranolazine, a novel antianginal agent with antiarrhythmic properties. *Circulation* **110**, 904–910 (2004).
- Antoons, G. *et al.* Late Na⁺ current inhibition by ranolazine reduces torsades de pointes in the chronic atrioventricular block dog model. *J. Am. Coll. Cardiol.* **55**, 801–809 (2010).
- Moretti, A. *et al.* Patient-specific induced pluripotent stem-cell models for long-QT syndrome. *N. Engl. J. Med.* **363**, 1397–1409 (2010).
- Huangfu, D. *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nature Biotechnol.* **26**, 795–797 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This study was supported in part by the Israel Science Foundation and Legacy Heritage Foundation (no. 1225/09), by the Yad Hanadiv Foundation Bruno Award, by the Lorry Lokey research fund, and by the Nancy and Stephen Grand Philanthropic Fund. We thank E. Suss-Toby and O. Shenker (from the multidisciplinary laboratory unit) and M. Tzukerman for their valuable help, A. Zamir for writing the MEA analysis software and I. Laevsky and T. Falik-Zaccai for the karyotype analysis.

Author Contributions I.I., L.M., I.H. and L.G. designed the experiments; I.I., L.M., I.H., L.Z.-D., O.C., A.W., O.F., A.G. and G.A. performed the experiments; I.I. and L.M. analysed and interpreted the electrophysiological data; M.B. and H.H. performed the clinical assessment; L.G. wrote the manuscript; all authors read and approved the manuscript; and L.G. supervised this research work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to L.G. (mdlior@technion.ac.il).

METHODS

Establishment and cardiomyocyte differentiation of patient-derived iPSCs.

Patient-derived human iPSC clones were established from the patient's skin fibroblasts by retroviral delivery of three reprogramming factors (SOX2, KLF4 and OCT4), followed by application of the histone deacetylase inhibitor valproic acid (VPA)²⁴. Briefly, Moloney-based retroviral vectors (pMXs) containing human complementary DNAs (cDNAs) of *OCT4*, *SOX2* and *KLF4* (Addgene plasmids 17964, 17218 and 17219, respectively) were used for retrovirus particle production. These plasmids were cotransfected with the helper plasmid encoding VSVG into HEK-293GP cells for virus production. Virus-containing media were collected at 48 and 72 h after transfection and used for two 24-h rounds of infection of the fibroblasts. Cells were then re-plated at a density of $1-2 \times 10^5$ cells per well on a MEF feeder layer, cultured in ES medium and treated with 0.9 mM VPA for 14 d.

Several human iPSC clones, which morphologically resembled human ES cells and were positively stained with vital TRA-1-81 or TRA-1-60 staining²⁵, were selected and expanded for further characterization. The undifferentiated human iPSC colonies were cultured on mitotically inactivated mouse embryonic fibroblast (MEF) cells as previously described^{11,13,26}. Y-27632 (5 μ M), a ROCK inhibitor, was added in early passages to increase seeding efficiency. The culture medium consisted of 80% knockout high-glucose glutamine-free DMEM with sodium pyruvate supplemented with 20% serum replacement, 1 mM L-glutamine, 0.1 mM mercaptoethanol, 4 ng ml⁻¹ human recombinant basic fibroblast growth factor (bFGF), and 1%-nonessential amino acid stock (Invitrogen).

Cardiomyocyte differentiation was induced using the embryoid body differentiating system as previously described in detail^{13,26}. Undifferentiated human iPSCs were dispersed into small cell clumps using 300 U ml⁻¹ of collagenase IV (Worthington) for 45 min. These cells were then cultivated in suspension for 10 days in the differentiation medium (knockout DMEM supplemented with 20% FBS (Hyclone), nonessential amino acids, β -mercaptoethanol and glutamine) where they aggregated to form embryoid bodies. The embryoid bodies were then plated on 0.1% gelatin-coated culture dishes, cultured with the same differentiation medium, and examined daily for the appearance of spontaneous contractions. At 30 days after plating, contracting areas within the embryoid bodies were carefully microdissected and used for the phenotypic characterization studies.

We studied cardiomyocytes derived from three different LQTS human iPSC clones (generated independently from the patient's fibroblasts). They were compared with cardiomyocytes derived from a healthy control human iPSC line generated using the same methods. We also used a second control human iPSC line that was well-characterized previously^{13,27}.

Genomic sequencing. Genomic DNA was isolated from the patient-derived human iPSC colonies using the high-pure PCR template preparation kit (Roche). The relevant DNA fragment of the *KCNH2* gene was amplified by PCR reaction using 100 ng genomic DNA (primer sequences are detailed in Supplementary Table 2). PCR products were then sequenced.

Immunofluorescence and alkaline phosphatase staining. Colonies of undifferentiated human iPSCs and single cardiomyocytes obtained by dissociating beating embryoid bodies were fixed with 4% paraformaldehyde, permeabilized with 1% Triton (Sigma) and blocked with 5% horse serum. Specimens were incubated overnight at 4 °C with primary antibodies targeting OCT4, Tra-1-60, CX43 (from Santa Cruz), NANOG (Peprotech), SSEA4 (R&D), nestin (Millipore), α -fetoprotein (Cell Marque), desmin (Thermo), troponin I (TnI; Chemicon) and sarcomeric α -actinin (Sigma). The preparations were incubated with secondary antibodies at a dilution of 1:200 for 1 h. Nuclei were counterstained with DAPI (1:500, Sigma). Preparations were examined using a laser-scanning confocal microscope (Zeiss LSM-510-PASCAL). Alkaline phosphatase activity was detected in live cultures using the alkaline phosphatase detection kit (Sigma) according to the manufacturer's instructions.

Teratoma formation. Undifferentiated human iPSCs were obtained by collagenase IV dissociation and injected subcutaneously into NOD/SCID mice. Palpable tumours were observed 4–6 weeks after injection. Tumour samples were collected at 6–8 weeks, cryo-preserved, processed by 10- μ m cryo-sectioning, and stained with haematoxylin and eosin.

Karyotype analysis. Karyotype analysis was performed using standard G-banding chromosome analysis by the institution's cytogenetic laboratory according to standard procedures.

Bisulphite sequencing. Genomic DNA (1 μ g) was bisulphite converted with the Methylamp DNA Modification kit (Epigenetec) according to the manufacturer's instructions. After bisulphite conversion, DNA was amplified using Faststart Taq polymerase (Roche) as follows: 94 °C for 4 min; 5 cycles of 94 °C for 30 s, 59 °C for 3 min, 72 °C for 3 min; 35 cycles of 94 °C for 30 s, 62 °C for 30 s, 72 °C for 30 s; 72 °C for 10 min. Primer sequences can be found in Supplementary Table 2. PCR products were TA-cloned into pTZ57R/T plasmid (Fermentas). Inserts were sequenced with M13 universal primers.

Gene expression analysis. Undifferentiated human iPSCs and differentiating embryoid bodies were frozen in liquid nitrogen. RNA was isolated using the RNeasy-plus mini-kit (Qiagen-Gm). Reverse transcription into cDNA was conducted using the high-capacity cDNA reverse transcription kit (Applied Biosystems). The PCR-related primers are detailed in Supplementary Table 2. Briefly, each RT-PCR included the following PCR program: 3 min at 93 °C, 30 s at 93 °C, 30 s at 60 °C, and 30 s at 72 °C. 2.5 ng of cDNA was used from each sample.

SYBR-Green real-time PCR studies were performed using the Fast SYBR Green Master mix (Applied Biosystems) and primers (detailed in Supplementary Table 3). All real-time PCR experiments were conducted in triplicates. Samples were cycled 40 times using a Fast ABI-7500 Sequence Detector (Applied Biosystems). ABI-7500 cycle conditions were as follows: 2 min at 50 °C, 15 min at 95 °C followed by 40 cycles of 15 s at 95 °C, 30 s at 60 °C and 30 s at 72 °C. Cycle threshold (CT) was calculated under default settings for real-time sequence detection software (Applied Biosystems).

Whole-cell patch-clamp recordings. Spontaneously contracting embryoid bodies were mechanically isolated, enzymatically dispersed into single cells (1 mg ml⁻¹ collagenase B; Roche) and attached to fibronectin-coated glass coverslips. Action potentials were recorded from spontaneously contracting small clusters superfused with Tyrode solution at 32 °C under spontaneous and paced (0.5–1 Hz) rhythms. The Tyrode solution consisted of NaCl (140 mM), KCl (5.4 mM), CaCl₂ (1.8 mM), MgCl₂ (1 mM), HEPES (10 mM) and glucose (10 mM), pH 7.4. The pipette solution consisted of KCl (120 mM), MgCl₂ (1 mM), Mg-ATP (3 mM), HEPES (10 mM) and EGTA (10 mM), pH, 7.2. Action potentials were recorded using the current-clamp mode using Axopatch 200B, Digidata 1322A, and pClamp 9 (Axon Instruments) for data amplification, acquisition and analysis, respectively. A custom-written computer program was used to measure APD₅₀, APD₇₀ and APD₉₀ (the time intervals required to reach 50%, 70% and 90% of repolarization), dV/dt_{max}, repolarization velocity and action-potential amplitude.

For voltage-clamp studies of I_{Kr} , external Na⁺ was replaced by equimolar choline (126 mM) and the solution was supplemented by 4-AP (5 mM), BaCl₂ (0.5 mM), CdCl₂ (0.2 mM) and chromanol (10 μ M) to suppress potential interference of I_{Na} , I_{to} , I_{K1} , I_{Ca} and I_{Ks} , respectively. I_{Kr} was defined as the E-4031-sensitive (1 μ M) current and was elicited by 3-s depolarizing steps from a holding potential of -60 mV to potentials ranging from -40 to +60 mV in 10 mV increments. This was followed by a 2-s repolarization phase to -30 mV to elicit tail current.

Microelectrode array recordings. A high-resolution microelectrode array recording system (Multichannel Systems)²⁸ was used to characterize the electrophysiological properties of human iPSC-CMs. The contracting areas within the embryoid bodies were microdissected and plated on fibronectin-coated microelectrode array plates. The recorded extracellular electrograms were used to determine local FPD, defined as the time interval between the initial deflection of the field potential to its return to baseline. FPD measurements were either compared in rate-matched specimens or were normalized (cFPD) to the activation rate using Bazett's, Hodges' or Fridericia's¹⁶ correction formulae. For Bazett's correction, cFPD-B = FPD/(RR)^{1/2}; for Hodges' correction, cFPD-H = FPD + 0.105 \times ((1/RR) - 1); for Fridericia's correction, cFPD-Fr = FPD/(RR)^{1/3}, where RR indicates the time interval (in seconds) between two consecutive beats.

To assess for the effects of different drugs on the electrophysiological properties of the cells being studied, between 5–30 μ l of tested drug stock solution was added to 1 ml of medium to achieve the necessary concentration. The tested drugs included isoproterenol hydrochloride, nifedipine, ranolazine dihydrochloride, pinacidil monohydrate (all from Sigma), cisapride (Janssen-Cilag) and E-4031 (Alomone Labs). Preliminary studies (data not shown) demonstrated that DMSO (solvent for nifedipine and pinacidil) alone did not have any effect on the measured electrophysiological parameters. Extracellular recordings were performed for 60 s at baseline and after 5 min following drug application.

Statistical analysis. Results are reported as mean \pm s.e.m. Comparison between LQTS human iPSC-CMs and control healthy human iPSC-CMs groups was performed using the unpaired Student's *t*-test. Drug effects on the cells were analysed by paired *t*-test. One-way ANOVA followed by Tukey's post-hoc tests were used when comparing multiple groups. *P* < 0.05 was considered statistically significant.

- Lowry, W. E. *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl Acad. Sci. USA* **105**, 2883–2888 (2008).
- Arbel, G. *et al.* Methods for human embryonic stem cells derived cardiomyocytes cultivation, genetic manipulation, and transplantation. *Methods Mol. Biol.* **660**, 85–95 (2010).
- Park, I. H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).
- Kehat, I., Gepstein, A., Spira, A., Itskovitz-Eldor, J. & Gepstein, L. High-resolution electrophysiological assessment of human embryonic stem cell-derived cardiomyocytes: a novel *in vitro* model for the study of conduction. *Circ. Res.* **91**, 659–661 (2002).

Using induced pluripotent stem cells to investigate cardiac phenotypes in Timothy syndrome

Masayuki Yazawa¹, Brian Hsueh¹†, Xiaolin Jia¹†, Anca M. Pasca¹†, Jonathan A. Bernstein², Joachim Hallmayer³ & Ricardo E. Dolmetsch¹

Individuals with congenital or acquired prolongation of the QT interval, or long QT syndrome (LQTS), are at risk of life-threatening ventricular arrhythmia^{1,2}. LQTS is commonly genetic in origin but can also be caused or exacerbated by environmental factors^{1,3}. A missense mutation in the L-type calcium channel *Ca_v1.2* leads to LQTS in patients with Timothy syndrome^{4,5}. To explore the effect of the Timothy syndrome mutation on the electrical activity and contraction of human cardiomyocytes, we reprogrammed human skin cells from Timothy syndrome patients to generate induced pluripotent stem cells, and differentiated these cells into cardiomyocytes. Electrophysiological recording and calcium (Ca^{2+}) imaging studies of these cells revealed irregular contraction, excess Ca^{2+} influx, prolonged action potentials, irregular electrical activity and abnormal calcium transients in ventricular-like cells. We found that roscovitine, a compound that increases the voltage-dependent inactivation of *Ca_v1.2* (refs 6–8), restored the electrical and Ca^{2+} signalling properties of cardiomyocytes from Timothy syndrome patients. This study provides new opportunities for studying the molecular and cellular mechanisms of cardiac arrhythmias in humans, and provides a robust assay for developing new drugs to treat these diseases.

The risk of sudden death due to genetic and drug-induced LQTS is a major concern for patients, clinicians and pharmaceutical companies. Genetic LQTS has an estimated prevalence of 1 in 7,000 individuals and results from mutations in at least 10 genes^{1,4,5,9–12}. Drug-induced LQTS is a side effect of many approved drugs and is a common cause of drug failure in clinical trials. Despite our knowledge of many of the genes that cause LQTS, the mechanisms that underlie the disease in humans are incompletely understood. Mouse models of human LQTS have proved to be problematic because the mouse resting heart rate is approximately tenfold faster than that of humans and therefore mouse cardiomyocytes have different electrical properties than their human counterparts. Therefore it is essential to develop models of LQTS that use human cardiomyocytes.

Ca_v1.2 is the main L-type channel in the mammalian heart and is essential for generating the cardiac action potential and for excitation contraction coupling^{13–15}. Ca^{2+} influx through L-type channels in the plasma membrane causes Ca^{2+} release through ryanodine receptors in the sarcoplasmic reticulum, leading to muscle contraction^{14–17}. A single amino acid substitution in exon 8a of *CACNA1C*, the gene encoding *Ca_v1.2* in humans, causes Timothy syndrome, a disorder characterized by LQTS, syndactyly (webbing of fingers and toes), immune deficiency and autism⁴. Exon 8a is an alternatively spliced exon of *Ca_v1.2*, and the Timothy syndrome mutation is a glycine to arginine (G to R) substitution that impairs inactivation of the channel^{4,5,18,19}. *Ca_v1.2* channels undergo both voltage-dependent inactivation and Ca^{2+} -dependent inactivation, and the G406R mutation severely impairs voltage-dependent inactivation and subtly affects Ca^{2+} -dependent inactivation. Precisely how this leads to LQTS or arrhythmias in humans is not known.

To generate induced pluripotent stem cells (iPSCs) from Timothy syndrome patients, we first obtained dermal fibroblasts from two patients by punch biopsy. We confirmed the presence of the Timothy syndrome mutation in these cells by sequencing the genomic DNA using two primer sets that recognize exon 8a in *CACNA1C* (ref. 4). We next reprogrammed the fibroblasts to generate iPSCs using four retroviruses containing *SOX2*, *OCT3/4* (also known as *POU5F1*), *KLF4* and *MYC* (refs 20, 21). Three to four weeks after the infection, we picked human embryonic stem cell (hESC)-like colonies based on their morphology and expanded them for characterization and *in vitro* differentiation into cardiac cells (Fig. 1a). We generated a total 16 iPSC lines from two Timothy syndrome patients, and 10 control lines from two unrelated individuals without Timothy syndrome.

We selected five Timothy syndrome and five control iPSC lines for further characterization and generation of cardiomyocytes (Supplementary Table 1). We used genomic sequencing to confirm that the Timothy syndrome iPSCs preserved the Timothy syndrome mutation (Supplementary Fig. 1), and we mapped the integration sites of the retroviruses using nested polymerase chain reaction (PCR)²² (Supplementary Table 1). We found that all the lines had independent retroviral insertion sites and that the sites did not occur in the coding region of any gene. To examine whether the iPSC lines express human embryonic stem cell markers, we used immunocytochemistry to look for *NANOG* and *TRA2-49-6E* (alkaline phosphatase) expression, and found that all of the lines expressed these markers (Supplementary Fig. 2). We also used reverse transcriptase PCR (RT-PCR) to confirm that iPSCs expressed *NANOG* and *Rex-1* and that they had silenced the exogenous genes that were used for reprogramming (Supplementary Fig. 3). In addition, we performed genome-wide microarray analyses of the cells and found that the gene expression patterns of iPSCs closely resembled those of hESCs and not of fibroblasts or neurons (GEO database <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=jncfzcucugewaha&acc=GSE25542>). We also karyotyped the iPSC lines to ensure that they did not have large chromosomal abnormalities (Fig. 1b and Supplementary Table 1) and injected them into immunodeficient mice to verify that they could generate teratomas. Both control and Timothy syndrome iPSCs were able to form tissues derived from all three germ layers including neural tissues (ectoderm), cartilage (mesoderm) and gut-like epithelium (endoderm) (Supplementary Fig. 2 and Supplementary Table 1), indicating that the iPSCs that we have generated are pluripotent.

To generate human cardiomyocytes from iPSCs, we first prepared embryoid bodies (EBs) from five control and five Timothy syndrome iPSC lines. After one week in suspension culture (day 7), we placed ~80 EBs onto gelatin-coated 100-mm dishes to allow EBs to attach. Thirty days (day 37) after plating, we observed that ~0.5–20% of the EBs showed rhythmic contractions (Fig. 1c). There were no significant differences in the number of contracting EBs formed from control and Timothy syndrome iPSC lines (data not shown). RT-PCR

¹Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA. ²Department of Pediatrics, Stanford University School of Medicine, Stanford, California 94305, USA.

³Department of Psychiatry & Behavioral Science, Stanford University School of Medicine, Stanford, California 94305, USA. †Present addresses: Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA (B.H.); Baylor College of Medicine, Houston, Texas 77030, USA (X.J.); Lucile Packard Children's Hospital, Stanford University, Palo Alto, California 94304, USA (A.M.P.).

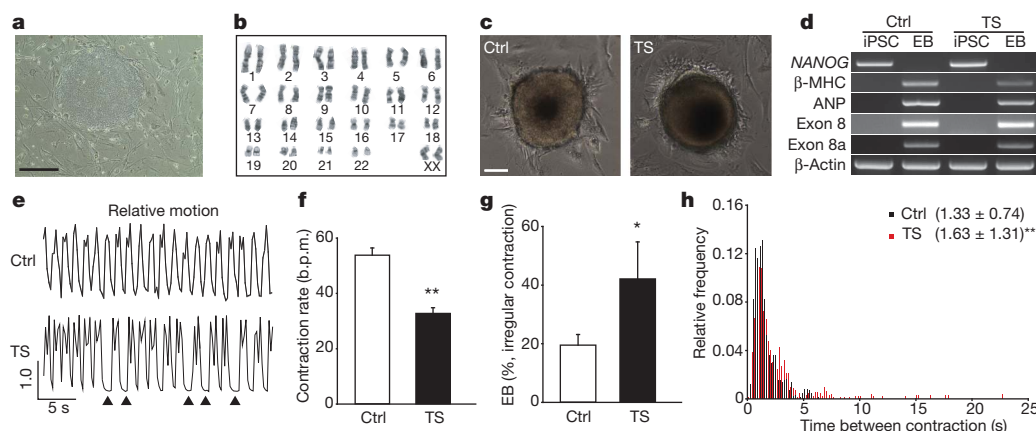


Figure 1 | Generation of cardiomyocytes from control and Timothy syndrome iPSCs. **a**, Phase contrast images of iPSC line (9862-61) derived from a Timothy syndrome patient. Scale bar, 400 μ m. **b**, Karyogram of Timothy syndrome iPSCs (7643-5). **c**, Images of spontaneously contracting embryoid bodies (EBs) generated from control (Ctrl, left) and Timothy syndrome iPSCs (TS, right). Scale bar, 100 μ m. **d**, Examination of pluripotent and cardiac gene expression using RT-PCR with primer sets for a pluripotent gene (*NANOG*), cardiac markers (β -MHC and ANP), $\text{Ca}_v1.2$ channels (exon 8 and 8a) and a housekeeping gene (β -actin). **e**, Relative motion of contracting control and Timothy syndrome EBs. Arrowheads show missing contractions.

analysis revealed that the spontaneously contracting EBs expressed cardiac markers (Fig. 1d) and both exons 8 and 8a of *CACNA1C* but not *NANOG*, indicating that these contain cardiomyocytes but not iPSCs.

To examine the contractile properties of cardiomyocytes we collected time-lapse images of spontaneously contracting EBs and analysed their movement using image analysis software (Supplementary Movies 1–4 and Fig. 1e, f). We collected movies of 113 EBs derived from five Timothy syndrome iPSC lines and compared them to EBs derived from the five control lines. Control EBs contracted at approximately 60 beats per minute (b.p.m.), similar to the resting heart rate in humans, whereas the Timothy syndrome cardiomyocytes contracted at only ~ 30 b.p.m. Contraction of the Timothy syndrome EBs was significantly more irregular than contraction of control cardiomyocytes (Fig. 1e, g). This was reflected in the broader distribution of inter-contraction intervals in Timothy syndrome EBs relative to those of controls (Fig. 1h). These results indicate that contracting EBs from multiple iPSC lines from two Timothy syndrome patients have disease-specific defects.

To characterize the underlying defects in cardiomyocytes from Timothy syndrome patients further, we dissociated contracting EBs into single cells. We stained these cells with antibodies that recognize the cardiac markers α -actinin, which is present at the Z-line of the sarcomere, and cardiac troponin I, which is a cardiac-specific myofibrillar protein (Fig. 2a and Supplementary Fig. 4). More than 65% of the cells from the beating EBs expressed both proteins and had well-organized sarcomeres. There were no differences in the staining pattern of Timothy syndrome and control cardiomyocytes. We next used whole-cell patch clamping to determine whether the Timothy syndrome mutation altered L-type channel voltage-dependent inactivation in human cardiomyocytes. We measured L-type-channel-generated currents using Ba^{2+} as the charge carrier and found that the L-type channel current in Timothy syndrome cardiomyocytes had significantly reduced voltage-dependent inactivation compared to control cells (Fig. 2b–e and Supplementary Fig. 4). This was apparent both from the increased amplitude of the residual current after a 350-ms depolarization and from increased current elicited by a 300-ms test pulse after a 2-s depolarizing pulse to different voltages. In contrast, there was no difference between control and Timothy syndrome

cardiomyocytes in the current–voltage relationship or the peak amplitudes of Ba^{2+} currents (Fig. 2d). These results are broadly consistent with the properties of the Timothy syndrome mutant channel observed in heterologous expression systems^{4,18,19}.

Ca^{2+} influx through L-type channels contributes to the plateau phase of the cardiac action potential, so we asked whether the shape or duration of the action potential in cardiomyocytes was altered by the Timothy syndrome mutation. Using current-clamp recording, we examined spontaneous action potentials in control and Timothy syndrome cardiomyocytes. Because human cardiomyocyte populations generated from iPSCs contain nodal-like, ventricular-like and atrial-like myocytes^{23,24}, we harvested the mRNA from each patch-clamped cardiomyocyte and used single-cell RT-PCR of cardiac ventricular myosin light chain 2v (*MLC2v*, Supplementary Fig. 5) to identify ventricular cells. We found that ventricular-like myocytes from Timothy syndrome patients had action potentials that were three times as long as those of control cells (Fig. 2f, g). In addition, the Timothy syndrome cardiomyocytes exhibited a large number of depolarizing events that failed to trigger a full action potential. These depolarizations were similar to the delayed afterdepolarizations (DADs) that arise after ectopic release of Ca^{2+} from the sarcoplasmic reticulum and which are associated with cardiac arrhythmias (Fig. 2f, h). In contrast we did not find significant differences in the action potential properties of nodal-like and atrial-like myocytes from control and Timothy syndrome patients (Supplementary Fig. 5). These findings indicate that ventricular cardiomyocytes derived from Timothy syndrome iPSCs are defective and provide a possible cellular basis for LQTS and arrhythmia in these patients.

$\text{Ca}_v1.2$ channels have a crucial role in activating Ca^{2+} -induced Ca^{2+} release from the sarcoplasmic reticulum but the effect of altering $\text{Ca}_v1.2$ inactivation on this process is not known. We therefore asked how the Timothy syndrome mutation affects Ca^{2+} signalling in Timothy syndrome cardiomyocytes. We used a confocal microscope to perform fast line-scan imaging of human cardiomyocytes loaded with the Ca^{2+} indicator Fluo-4. The Ca^{2+} elevations in spontaneously contracting Timothy syndrome cardiomyocytes were more irregular than those of control cardiomyocytes (Fig. 3 and Supplementary Fig. 6). In addition, the Timothy syndrome mutation led to significantly larger and more prolonged Ca^{2+} elevations, indicating that channel

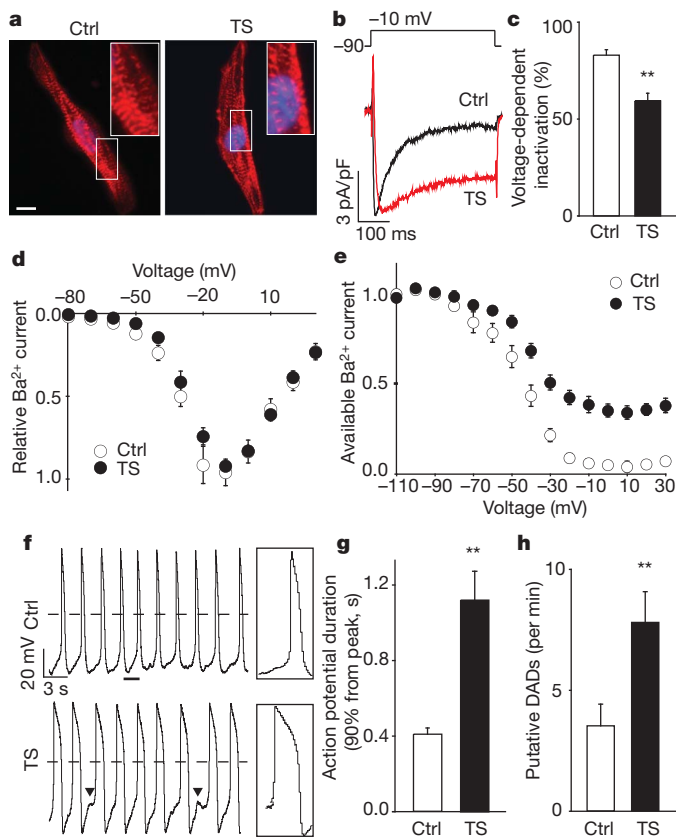


Figure 2 | Electrophysiological features of Timothy syndrome cardiomyocytes. **a**, Immunocytochemistry of human cardiomyocytes generated from control (left) and Timothy syndrome iPSCs (right) using anti- α -actinin antibodies (red). The nuclei (blue) are marked by Hoechst staining. Insets show high magnification images of the sarcomeres. Scale bar, 10 μm . **b**, Voltage-clamp recording of Ba^{2+} currents in control (black) and Timothy syndrome (red) cardiomyocytes show a defect in voltage-dependent channel inactivation after a voltage pulse from -90 to -10 mV. **c**, Voltage-dependent inactivation in control and Timothy syndrome cardiomyocytes 350 ms after the start of the pulse (** $P < 0.01$; Student's t -test). **d**, The I - V relationship of Timothy syndrome (filled circles) and control (open circles) Ca^{2+} currents (mean \pm s.e.m.) are statistically identical. There were no significant differences in the peak amplitude of Ba^{2+} currents between control and Timothy syndrome cardiomyocytes (data not shown). **e**, Ba^{2+} current in control and Timothy syndrome myocytes stimulated with a test pulse to $+10$ mV after a family of prepulses from -110 to $+30$ mV in 10 -mV increments (raw traces in Supplementary Fig. 4c; control, $n = 23$ cells in 4 lines; Timothy syndrome, $n = 19$ in 4 lines, mean \pm s.e.m.). **f**, Spontaneous action potentials in control and Timothy syndrome ventricular-like myocytes measured in current-clamp mode. Boxes show the regions indicated by underlines at an expanded timescale. Arrowheads show putative delayed afterdepolarizations (DADs). Dashed lines show 0 mV. **g**, Action potential duration in Timothy syndrome and control ventricular cardiomyocytes. The expression of the ventricular marker, MLC2v , was confirmed with single-cell RT-PCR immediately after whole-cell patch recording (Supplementary Fig. 5). **h**, Putative DADs in Timothy syndrome and control ventricular-like cardiomyocytes (control, $n = 22$ cells in 4 lines; Timothy syndrome, $n = 14$ in 4 lines, mean \pm s.e.m.). Statistical analyses were performed with Student's t -test (** $P < 0.01$).

inactivation is important for maintaining the timing and the amplitude of the ventricular Ca^{2+} release.

The finding that cardiomyocytes from Timothy syndrome patients have disease-specific electrical defects indicates that they might be a useful system for testing possible therapeutic compounds. As a proof of principle, we investigated whether roscovitine, a cyclin-dependent kinase inhibitor that increases voltage-dependent inactivation in HEK 293 cells expressing $\text{Ca}_v1.2$ channels^{6–8}, could rescue the phenotypes of Timothy syndrome cardiomyocytes. We examined the effect of

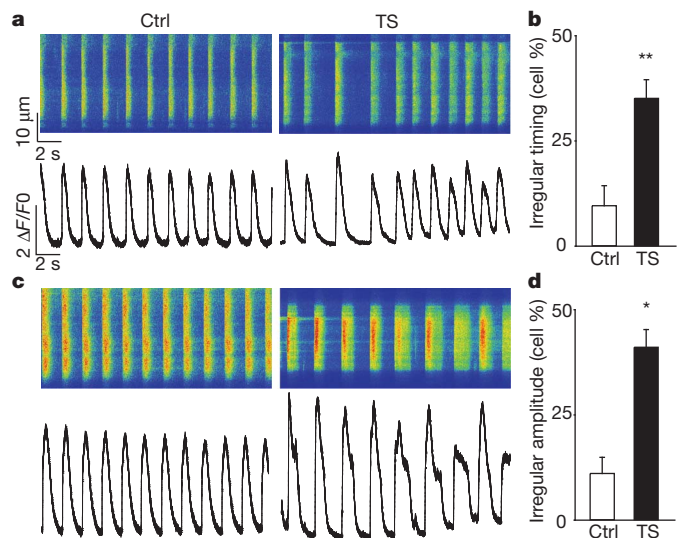


Figure 3 | Ca^{2+} signalling in Timothy syndrome and control cardiomyocytes. **a**, **c**, Representative line-scan images (top) and spontaneous Ca^{2+} transients (bottom) in control (left) and Timothy syndrome cardiomyocytes (right). **b**, **d**, Timothy syndrome cardiomyocytes showed more irregular timing (**b**) and amplitude (**d**) of spontaneous Ca^{2+} transients compared to control cells (see Supplementary Fig. 6 and Methods for details about the analysis; control, $n = 102$ cells in 4 lines; Timothy syndrome, $n = 149$ in 4 lines, mean \pm s.e.m.). Statistical analyses were performed with Student's t -test (* $P < 0.05$, ** $P < 0.01$).

three different concentrations of roscovitine (10, 33.3 and 100 μM) on the timing and amplitude of spontaneous Ca^{2+} transients in Timothy syndrome cardiomyocytes. Treatment with 100 μM roscovitine completely eliminated contractions but 33.3 μM roscovitine significantly reduced both the irregular timing and amplitude of Ca^{2+} transients (Fig. 4a, b and Supplementary Fig. 7). Washing out roscovitine partially restored the irregular Ca^{2+} transients in Timothy syndrome cardiomyocytes but this effect did not reach significance^{6–8}.

To determine if roscovitine rescues the electrophysiological properties of Timothy syndrome cardiomyocytes, we used whole-cell patch clamping to measure $\text{Ca}_v1.2$ currents and action potentials in Timothy syndrome cardiomyocytes in the presence and absence of roscovitine. Roscovitine significantly increased $\text{Ca}_v1.2$ voltage-dependent inactivation in Timothy syndrome cardiomyocytes (Fig. 4c, d) but had only a mild effect on control cardiomyocytes (Supplementary Fig. 8). Roscovitine also reduced the duration of action potentials in Timothy syndrome cardiomyocytes and decreased the frequency of abnormal depolarizing events (Fig. 4e, f). These results indicate that cardiomyocytes from patients with LQTS can be used to screen potential drugs, and suggest that drugs related to roscovitine might be valuable tools for treating Timothy syndrome and other cardiac arrhythmias.

We have developed a new *in vitro* model for studying cardiac arrhythmias that has important advantages relative to existing approaches²⁵. In contrast to mouse cardiomyocytes, iPSC-derived EBs spontaneously contract at a rate that is similar to that of the human heart, and single cells derived from these EBs have structural and electrical properties that are similar to those of cardiomyocytes from human patients. Importantly, cardiomyocytes derived from five different iPSC lines from two independent patients with LQTS had cellular defects that are consistent with the cardiac defects of the patients. The EBs from Timothy syndrome patients contracted slowly relative to control EBs, consistent with bradycardia in many Timothy syndrome patients. Ventricular cardiomyocytes from Timothy syndrome patients also had prolonged action potentials that probably delay the repolarization of the heart and lead to LQTS. Both the isolated cardiomyocytes and the EBs contracted arrhythmically and the cardiomyocytes had

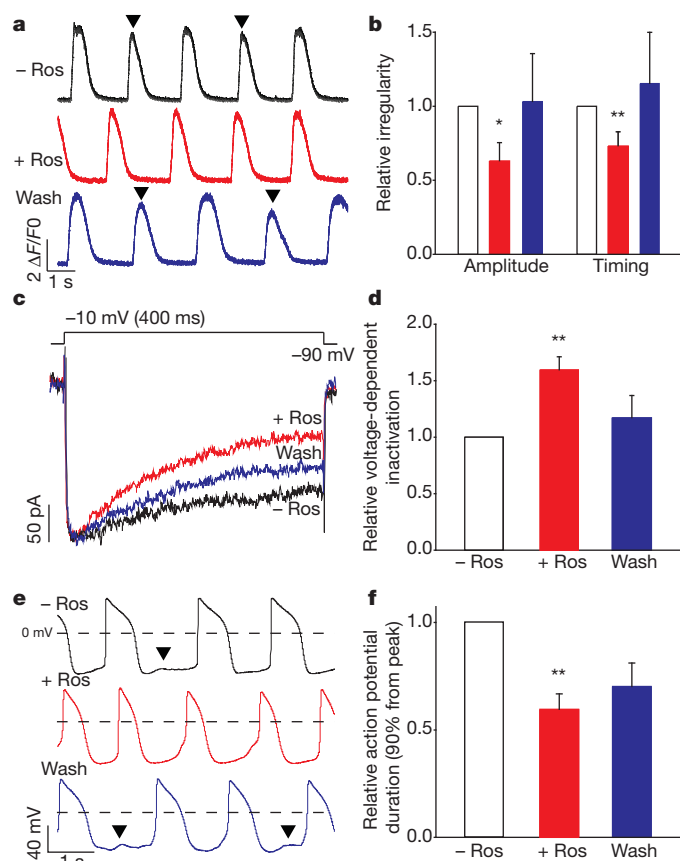


Figure 4 | Roscovitine rescues the cellular phenotypes of Timothy syndrome cardiomyocytes. **a**, Spontaneous Ca^{2+} transients in Timothy syndrome cardiomyocytes before (black) and during treatment with 33.3 μM roscovitine (red) as well as after wash-out (blue). Arrowheads show irregular Ca^{2+} elevations. **b**, Effects of roscovitine on the relative irregularity of the amplitude and timing of the spontaneous Ca^{2+} transients in Timothy syndrome cardiomyocytes ($n = 8$ cells in 2 lines, * $P < 0.05$, ** $P < 0.01$, mean \pm s.e.m.). **c**, Ba^{2+} currents in Timothy syndrome cardiomyocytes recorded in voltage-clamp mode before (black) during (red) and after (blue) treatment with 33.3 μM roscovitine. Roscovitine promoted inactivation of currents in Timothy syndrome cardiomyocytes. **d**, Effects of roscovitine on $\text{Ca}_v1.2$ voltage-dependent inactivation in cardiomyocytes ($n = 5$ cells in 2 lines, ** $P < 0.01$, mean \pm s.e.m.). **e**, Spontaneous action potentials recorded in current-clamp recording before, during and after treatment with roscovitine. Arrowheads show putative DADs. **f**, Roscovitine prevented action potential prolongation observed in Timothy syndrome cardiomyocytes ($n = 8$ cells in 2 lines, ** $P < 0.01$, mean \pm s.e.m.).

frequent depolarizing events that failed to produce action potentials and were similar to DADs.

The phenotype of Timothy syndrome cardiomyocytes stands in contrast to the phenotype of cardiomyocytes from patients with LQTS1 (ref. 12). Only ventricular cardiomyocytes from Timothy syndrome patients had prolonged action potentials, whereas both ventricular- and atrial-like cardiomyocytes from LQTS1 had this phenotype. Furthermore, arrhythmias and delayed depolarizations were observed in spontaneously beating Timothy syndrome cardiomyocytes whereas they could only be elicited in LQTS1 cardiomyocytes by stimulation with isoproterenol. Although it is difficult to link these features to the torsade de points and to ventricular fibrillations in Timothy syndrome patients, these findings set the stage for the development of more sophisticated models of LQTS. Finally, this study demonstrates that iPSC-derived cardiomyocytes are a useful platform for identifying drug candidates. Roscovitine restored the action potential duration in Timothy syndrome cardiomyocytes and prevented the occurrence of arrhythmias. Even though roscovitine has

other targets²⁶ it could be a useful lead compound for the development of new types of antiarrhythmics.

METHODS SUMMARY

Control and Timothy syndrome iPSC lines were generated using retroviral infection with pMXs-SOX2, pMXs-OCT3/4, pMXs-MYC and pMXs-KLF4 expression plasmids (Addgene) generated by S. Yamanaka's group²⁰. The iPSCs were cultured on irradiated DR4 mouse embryonic fibroblast feeders using standard ES media with 10–15 ng ml⁻¹ bFGF (R & D Systems), and cells were passaged with dispase (3 unit ml⁻¹, Invitrogen). The G1216A substitution in exon 8a was detected by sequencing of PCR products from DNA harvested from fibroblasts and iPSC lines using primers for human $\text{Ca}_v1.2$ exon 8a. Immunocytochemistry, RT-PCR, microarray, karyotyping and teratoma formation assay were performed using standard protocols. For *in vitro* generation of cardiomyocytes, embryoid bodies were cultured with Wnt3a (100 ng ml⁻¹, R&D Systems)²⁷. Whole-cell patch clamp recordings in single cardiomyocytes were conducted using standard methods. Live-cell Ca^{2+} imaging was performed in single cardiomyocytes loaded with 5 μM Fluo-4 AM and 0.02% Pluronic F-127 (Molecular Probes) using fast line scanning (1.92 ms/line) on a confocal microscope (LSM 510 Meta, Carl Zeiss) with a $\times 63$ lens (NA = 1.4). R-roscovitine was obtained from Sigma-Aldrich.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 November 2010; accepted 21 January 2011.

Published online 9 February 2011.

- Keating, M. T. The long QT syndrome. A review of recent molecular genetic and physiologic discoveries. *Medicine* **75**, 1–5 (1996).
- Huikuri, H. V., Castellanos, A. & Myerburg, R. J. Sudden death due to cardiac arrhythmias. *N. Engl. J. Med.* **345**, 1473–1482 (2001).
- Paakkari, I. Cardiotoxicity of new antihistamines and cisapride. *Toxicol. Lett.* **127**, 279–284 (2002).
- Splawski, I. et al. $\text{Ca}_v1.2$ Ca^{2+} channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell* **119**, 19–31 (2004).
- Splawski, I. et al. Severe arrhythmic disorder caused by cardiac L-type Ca^{2+} channel mutations. *Proc. Natl Acad. Sci. USA* **102**, 8089–8096 (2005).
- Yarotsky, V. & Elmslie, K. S. Roscovitine, a cyclin-dependent kinase inhibitor, affects several gating mechanisms to inhibit cardiac L-type ($\text{Ca}_v1.2$) Ca^{2+} channels. *Br. J. Pharmacol.* **152**, 386–395 (2007).
- Yarotsky, V., Gao, G., Peterson, B. Z. & Elmslie, K. S. The Timothy syndrome mutation of cardiac $\text{Ca}_v1.2$ (L-type) channels: multiple altered gating mechanisms and pharmacological restoration of inactivation. *J. Physiol. (Lond.)* **587**, 551–565 (2009).
- Yarotsky, V. et al. Roscovitine binds to novel L-channel ($\text{Ca}_v1.2$) sites that separately affect activation and inactivation. *J. Biol. Chem.* **285**, 43–53 (2010).
- Roden, D. M. & Viswanathan, P. C. Genetics of acquired long QT syndrome. *J. Clin. Invest.* **115**, 2025–2032 (2005).
- Chen, L. et al. Mutation of an A-kinase-anchoring protein causes long-QT syndrome. *Proc. Natl Acad. Sci. USA* **104**, 20990–20995 (2007).
- Roden, D. M. Clinical practice. Long-QT syndrome. *N. Engl. J. Med.* **358**, 169–176 (2008).
- Moretti, A. et al. Patient-specific induced pluripotent stem-cell models for long-QT syndrome. *N. Engl. J. Med.* **363**, 1397–1409 (2010).
- Reuter, H. Ion channels in cardiac cell membranes. *Annu. Rev. Physiol.* **46**, 473–484 (1984).
- Flucher, B. E. & Franzini-Armstrong, C. Formation of junctions involved in excitation-contraction coupling in skeletal and cardiac muscle. *Proc. Natl Acad. Sci. USA* **93**, 8101–8106 (1996).
- Seisenberger, C. et al. Functional embryonic cardiomyocytes after disruption of the L-type $\alpha_1\text{C}$ ($\text{Ca}_v1.2$) Ca^{2+} channel gene in the mouse. *J. Biol. Chem.* **275**, 39193–39199 (2000).
- Takeshima, H. et al. Embryonic lethality and abnormal cardiac myocytes in mice lacking ryanodine receptor type 2. *EMBO J.* **17**, 3309–3316 (1998).
- Yazawa, M. et al. TRIC channels are essential for Ca^{2+} handling in intracellular stores. *Nature* **448**, 78–82 (2007).
- Barrett, C. F. & Tsien, R. W. The Timothy syndrome mutation differentially affects voltage- and calcium-dependent inactivation of $\text{Ca}_v1.2$ L-type Ca^{2+} channels. *Proc. Natl Acad. Sci. USA* **105**, 2157–2162 (2008).
- Thiel, W. H. et al. Proarrhythmic defects in Timothy syndrome require calmodulin kinase II. *Circulation* **118**, 2225–2234 (2008).
- Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
- Yu, J. et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
- Aoi, T. et al. Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* **321**, 699–702 (2008).
- He, J. Q., Ma, Y., Lee, Y., Thomson, J. A. & Kamp, T. J. Human embryonic stem cells develop into multiple types of cardiac myocytes: action potential characterization. *Circ. Res.* **93**, 32–39 (2003).
- Zhang, J. et al. Functional cardiomyocytes derived from human induced pluripotent stem cells. *Circ. Res.* **104**, e30–e41 (2009).

25. Brunner, M. *et al.* Mechanisms of cardiac arrhythmias and sudden death in transgenic rabbits with long QT syndrome. *J. Clin. Invest.* **118**, 2246–2259 (2008).
26. Meijer, L. *et al.* Biochemical and cellular effects of roscovitine, a potent and selective inhibitor of the cyclin-dependent kinases cdc2, cdk2 and cdk5. *Eur. J. Biochem.* **243**, 527–536 (1997).
27. Tran, T. H. *et al.* Wnt3a-induced mesoderm formation and cardiomyogenesis in human embryonic stem cells. *Stem Cells* **27**, 1869–1878 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Timothy and the Timothy syndrome patients who participated in this study; U. Francke for discussion and for providing karyotyping expertise; A. Cherry and D. Bangs for fibroblast isolation; K. C. Chan for iPSC cultures; O. Shcheglovitov for help with electrophysiological recordings; and A. Olson for help with the confocal microscope. Funding was provided by grants from the Japan Society

for the Promotion for Science and the American Heart Association Western States to M.Y., and a National Institutes of Health Director's Pioneer Award, a grant from the Simons Foundation to R.E.D and gifts from L. Miller, B. and F. Horowitz and M. McCaffery.

Author Contributions M.Y. and R.E.D. designed research and wrote the manuscript; J.A.B., J.H. and R.E.D. recruited the Timothy syndrome patients; M.Y. and X.J. generated and characterized control and Timothy syndrome iPSCs; A.M.P. conducted karyotyping; M.Y. performed generation and characterization of human cardiomyocytes, whole-cell patch clamp, and Ca^{2+} imaging; M.Y. and B.H. analysed cardiomyocytes contraction rates.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.E.D. (ricardo.dolmetsch@stanford.edu).

METHODS

Cell culture. HEK 293T cells were cultured in Dulbecco's Modified Eagle Media (DMEM) supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, 100 unit ml⁻¹ penicillin and 100 µg ml⁻¹ of streptomycin (all from Invitrogen). Human fibroblasts were maintained in Minimum Essential Medium, Eagle (MEME, ATCC) containing 10% FBS, 100 unit ml⁻¹ penicillin and 100 µg ml⁻¹ of streptomycin. The patient fibroblasts were obtained by punch biopsy following a protocol approved by the Stanford Institutional Review Board (IRB). Human iPSCs were cultured on irradiated DR4 mouse embryonic fibroblast feeders using DMEM/F12 (1:1) medium containing 20% knockout SR, 1 mM non-essential amino acids (all from Invitrogen), 3 mM L-glutamine, 0.1 mM β-mercaptoethanol (Sigma-Aldrich), 100 unit ml⁻¹ penicillin and 100 µg ml⁻¹ streptomycin and 10–15 ng ml⁻¹ bFGF (R & D Systems). Human cardiomyocytes generated from iPSCs were maintained in EB5 media: DMEM/F12 with GlutaMAX (Invitrogen) containing 5% FBS, 1 mM non-essential amino acids, 1 mM L-glutamine, 0.1 mM β-mercaptoethanol, 100 unit ml⁻¹ penicillin and 100 µg ml⁻¹ streptomycin.

Plasmids. pMXs-SOX2, pMXs-OCT3/4, pMXs-MYC and pMXs-KLF4 expression plasmids were generated by S. Yamanaka's group²⁰ and were obtained from Addgene. The pUMVC packaging plasmid and pCMV-VSV-G envelope plasmids were used to prepare the retroviruses. A retrovirus encoding yellow fluorescent protein (YFP) was generated to examine the infection efficiency of human fibroblasts. pMXs-YFP was amplified using Taq polymerase, subcloned into the vector fragment of pMXs-OCT3/4, digested with EcoRI and verified by sequencing.

iPSC generation with retroviral infection. To prepare the retroviruses, HEK 293T cells were plated at 6–8 × 10⁵ cells per well in 6-well plates coated with 20 µg ml⁻¹ of poly-ornithine. Twenty-four hours after plating, the cells were transfected with pMXs plasmids containing the coding sequences for YFP, SOX2, OCT3/4, MYC or KLF4, along with pUMVC and pCMV-VSV-G plasmids at the ratio of 2.125 µg of plasmid DNA to 5 µl of Lipofectamine2000 in 500 µl OpiMEM (both from Invitrogen). The ratio of pMXs:pUMVC:pCMV-VSV-G was 8:8:1. Twenty-four hours after transfection, the DMEM media was replaced with fibroblast (MEME) media. The next day the culture media containing the retroviruses was harvested and mixed 1:1 with 0.6 µg ml⁻¹ polybrene (Sigma-Aldrich). The cell debris in the mixtures was removed using 0.45 µm low-protein-binding filters (Nalgene).

The patient and human fibroblasts were plated at 1 × 10⁵ cells per well in 6-well plates one day before infection with the retroviruses. The retrovirus solution was added to the fibroblasts for 8 h per day for 4 days to reduce cell death. The viruses were prepared fresh for each infection. To prevent cell crowding which can inhibit cell growth, the infected fibroblasts were passaged every other day with a dilution ratio of 1:3. Six days after the first infection, infected fibroblasts were plated at 1–50 × 10³ cells per dish in gelatin-coated 100 mm dishes containing ~1.5 × 10⁶ irradiated DR4 feeder cells per dish. The density of infected cells on feeders was critical for successful reprogramming so cells were always seeded at three different densities. One day after plating on DR4 feeder cells, the media was changed to human iPSC media made as described above. Human ES-like colonies were observed 21–30 days after the initial infection. These colonies were treated with 3 unit ml⁻¹ dispase (Invitrogen) for 10 min at 37 °C and washed three times with iPSC media, before being mechanically isolated with 22G needles under a microscope for subcloning.

iPSC characterization. The Timothy syndrome mutation in patient-derived cells was confirmed using DNA sequencing. Genomic DNA was isolated using DNeasy Blood and Tissue Kit (Qiagen) and the region encoding CACNA1C exon 8a was amplified using PCR and the primer set: forward 5'-TACACTAATCATCATAGGTCAT-3'; reverse 5'-TAGCGATTCCAGTTTAGGTAC-3'. The PCR product (~1.2 kb) was purified using the QIAquick PCR purification kit (Qiagen) and directly sequenced using the original forward primer and a new reverse primer (Ori-R-MY): 5'-CTCAGAGATAGACTGCTCAGTCTATG-3'.

To investigate the number and site of viral integration in iPSC genome we used nested PCR to amplify iPSC genomic DNA treated with BamHI + BglII or PstI as has been previously described²². The primer sets used are shown in Supplementary Table 2. The PCR products were purified and direct-sequenced or subcloned into pCR4-TOPO vector (Invitrogen).

To examine the expression of pluripotent markers, we placed the iPSCs on 15-mm round coverslips (Warner Instruments, CS-15R) coated with Matrigel (BD Biosciences) in 24-well plates. The iPSCs were fixed using 4% paraformaldehyde (PFA, Electron Microscopy Sciences, EM grade), 2% sucrose in PBS for 20 min at room temperature and washed three times with PBS. The cells were then blocked overnight with a solution containing 5% donkey serum (Calbiochem), 1% bovine serum albumin (BSA, Sigma-Aldrich) and 0.1% Triton X-100 (Sigma-Aldrich) in PBS for anti-NANOG staining or with a solution containing 3% BSA and 0.25% Triton X-100 in PBS for anti-TRA-2-49/6E antibody staining. The samples were then washed six times using PBS and incubated with the anti-NANOG (1/20 dilution in the

NANOG blocking solution; R&D Systems) and TRA-2-49/6E antibodies (1/80 dilution in the blocking solution without any detergents; the antibody developed by P. W. Andrews was obtained from Development Studies Hybridoma Bank developed under the auspices of the NICHD and maintained in the University of Iowa, Department of Biological Sciences) for at least 24 h at 4 °C. After incubation with primary antibodies, samples were washed six times (5 min each), and incubated in secondary antibody, Alexa Fluor 594 (Molecular Probes, 2 µg ml⁻¹), for 30 min at room temperature, washed with PBS six times (5 min each), stained with Hoechst 33285 (Molecular Probes, 1/10,000 dilution in PBS) and mounted on slides (Becton, Dickinson) using Aqua Poly/Mount (Polysciences). Images were acquired using a ×10 (NA = 0.5) objective lens or ×40 (NA = 1.3) oil immersion objective lens (Nikon) on the stage of a Nikon Eclipse TE2000U inverted microscope.

Karyotyping of iPSCs was achieved using standard protocols. For RT-PCR, RNA from fibroblasts and iPSCs was prepared using the RNeasy Mini kit and RNase-Free DNase set (Qiagen). cDNA was synthesized from 1 µg RNA using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen). The cDNA (21 µl) was diluted with DNase-free water (Invitrogen) at 1:5 and 1 µl of the samples was used for conventional RT-PCR with AccuPower PCR PreMix (Bioneer, K-2016) using primer sets (Supplementary Table2). For quantitative RT-PCR, FastStart Universal SYBR Green Master (Rox, Roche) and Mastercycler ep realplex (Eppendorf) were used.

Teratoma formation assays. Teratoma formation experiments were performed by injecting subcutaneously ~1 × 10⁶ cells in 50% Matrigel and 10 µM Y-27632 (Calbiochem) in iPSC media (~1 × 10⁷ cell ml⁻¹) into 8-week-old severe combined immunodeficient (SCID) beige mice (Charles River Laboratories). The mice were killed ~8 weeks after injection, the tumours were dissected, fixed with PBS containing 4% PFA and 10–30% sucrose and embedded in Tissue-Tek OCT compound (Sakura Finetek). Sections were stained with Haematoxylin 7211 and Eosin-Y (Thermo Scientific). All animal protocols and handling for the teratoma formation assay were performed following the guidelines established by Stanford University Administrative Panel on Laboratory Animal Care (APLAC) Committee.

Differentiation of iPSCs into cardiomyocytes. iPSCs in 6-well plates (Falcon, 353046) were treated with 3 unit ml⁻¹ of dispase for ~20 min at 37 °C to detach the colonies completely. The colonies were collected by allowing the iPSCs colonies settle to the bottom of the 50 ml tubes for ~6 min. The iPSC colonies were then washed and placed in Ultra-low 100-mm dishes (Corning Life Sciences) with iPSC media containing 10 µM Y-27632 without bFGF to form embryoid bodies. The next day the media was changed to EB20 with Wnt3a (DMEM/F12 with GlutaMAX containing 20% FBS, 1 mM non-essential amino acids, 1 mM L-glutamine, 0.1 mM β-mercaptoethanol, 100 unit ml⁻¹ penicillin, 100 µg ml⁻¹ streptomycin and 100 ng ml⁻¹ Wnt3a (R&D Systems²⁷)). Between days three and six, the media was changed to EB20 media without Wnt3a and media replaced every other day. At day 7, the EBs (~80 per dish) were plated into 100-mm dishes (BD Biosciences) coated with gelatin in EB5 media as described above. During days 7–37, the media was changed daily, and at around day 17 some of the EBs were observed to contract spontaneously. At day 37, EBs were isolated mechanically using a 22G needle, placed into 15-ml tubes, washed with PBS once, centrifuged (1,000 r.p.m., 2 min) and incubated in 0.25% trypsin and 0.03% EDTA (Invitrogen) for 4 min at room temperature with gentle shaking. EB5 media was then added to the tube, the cells were isolated by centrifugation and washed before using 1 ml of EB5 media and a P1000 filter tip (~30 times) to dissociate the EBs into single cells. The cells were plated on gelatin-coated 15-mm round coverslips in 24-well plates or 4-well LAB-TEK II chambers (Nalge Nunc International, chamber #1.5 German coverglass system). A single EB was sufficient for a few coverslips or wells of a chamber. For about 7 days after dissociation, cells were observed without conducting experiments, to ensure that the coverslips and chambers were not contaminated with fast-growing cell populations such as fibroblast-like cells and progenitors. The antibodies to cardiac Troponin I (clone 284, 1/200 dilution, Abcam) and α-actinin (clone EA-53, 1/200 dilution, Sigma-Aldrich) were used to characterize the dissociated cardiomyocytes.

Analysis of cardiomyocytes contraction. Images were collected at a rate of 5 frames per second using OpenLab software (Perkin Elmer) and converted to multiframe TIFF images for analysis using the Image Processing Toolbox in MATLAB R2009b (Mathworks). Relative motion between successive frames was quantified by subtracting each frame from the preceding frame and summing across all pixels. Movement was calculated by plotting the relative motion over time. Contractions were detected as a peak of relative motion and a second, typically smaller peak, corresponded to the relaxation of the EBs. Peaks were selected manually and the length of time between contractions was measured. Rhythmicity was measured by calculating the ratio of the standard deviation to the mean of the intervals between contractions. The threshold for classifying an EB

as contracting arrhythmically was two standard deviations from the average rhythmicity of control cells.

Patch-clamp electrophysiology. Whole-cell patch-clamp recordings of human cardiomyocytes generated from iPSCs were conducted using an EPC-10 patch-clamp amplifier (HEKA) and an inverted microscope equipped with differential interface optics (Nikon, TE2000-U). The glass pipettes were prepared using borosilicate glass (Sutter Instrument, BF150-110-10) using a micropipette puller (Sutter Instrument, Model P-87). Voltage-clamp measurements were conducted using an extracellular solution consisting of 5 mM BaCl₂, 160 mM TEA and 10 mM HEPES (pH 7.4 at 25 °C) and a pipette solution of 125 mM CsCl, 0.1 mM CaCl₂, 10 mM EGTA, 1 mM MgCl₂, 4 mM MgATP, 0.3 mM NaGTP and 10 mM HEPES (pH 7.4 with CsOH at 25 °C). The following pulse protocols were used: cells were held at -90 mV and depolarized to +30 mV for 400 ms at a rate of 0.1 Hz (Fig. 2b-d and Supplementary Fig. 4b); cells were held at -90, stimulated with a 2-s family of pulses from -110 to +30 which was followed by a 300-ms test pulse to -10 mV (Fig. 2e and Supplementary Fig. 4c). Current-clamp recording were conducted in normal Tyrode solution containing 140 mM NaCl, 5.4 mM KCl, 1 mM MgCl₂, 10 mM glucose, 1.8 mM CaCl₂ and 10 mM HEPES (pH 7.4 with NaOH at 25 °C) using the pipette solution: 120 mM K D-gluconate, 25 mM KCl, 4 mM MgATP, 2 mM NaGTP, 4 mM Na₂-phospho-creatin, 10 mM EGTA, 1 mM CaCl₂ and 10 mM HEPES (pH 7.4 with KCl at 25 °C). *R*-roscovitine (Sigma-Aldrich, R7772) was dissolved in DMSO and the same concentration of DMSO was used as a control.

Single-cell RT-PCR. Immediately after whole-cell patch-clamp recording, the glass recording pipette was changed to a larger glass pipette containing ~5 µl of lysis buffer (CellsDirect one-sep for qRT-PCR, Invitrogen) to harvest the intact cell for cDNA preparation. Harvesting pipettes were prepared using borosilicate glass with filament (Warner Instruments, G150T-4) with the following parameters: (heat, velocity, time) (1) 730, 20, 250; (2) 720, 20, 250; (3) 720, 20, 250; (4) 640,

36, 250, using a micropipette puller. The lysis solution containing the cell was transferred to a sterile PCR tube containing ~5 µl of the lysis buffer (total volume ~8 µl) using pressure from a syringe. The sample was incubated at 75 °C for 5 min and then chilled to 4 °C. Next, 10 µl of 2× RT Reaction Mix and 2 µl of RT Enzyme Mix (SuperScript III First-Strand Synthesis SuperMix for qRT-PCR, Invitrogen) were added. To synthesize cDNA from single cells the sample was incubated at 25 °C for 10 min; 50 °C for 30 min; 85 °C for 5 min; 4 °C for 5 min. To remove RNA, 1 µl of RNase H was added and the sample was incubated at 37 °C for 20 min. To perform RT-PCR, 1 µl of the samples were used as template for a 60–70 cycles.

Ca²⁺ imaging. Cardiomyocytes were placed in 4-well slide chambers and loaded with 5 µM Fluo-4 AM and 0.02% Pluronic F-127 (both reagents from Molecular Probes) in Tyrodes solution for 15 min at 37 °C. Ca²⁺ imaging was conducted with a confocal microscope (Carl Zeiss, LSM 510 Meta) with a ×63 lens (NA = 1.4) operated by Zen software (Carl Zeiss). Line scans were acquired at a sampling rate of 1.92 ms line⁻¹ (total 10,000 times for 19.2 s recording) at room temperature. Fiji software, a derivative of ImageJ (National Institutes of Health), Excel (Microsoft) and programs developed by B.H. were used to analyse irregularity of spontaneous Ca²⁺ transients.

Analysis of calcium imaging traces. The time course of intracellular Ca²⁺ was determined by using Fiji software to average the intensity across each line. The Ca²⁺ traces were then analysed using MATLAB to calculate the irregularity of the timing and amplitude of the Ca²⁺ transients. The period of the transients (timing) was defined as the peak-to-peak interval for two successive Ca²⁺ elevations with peaks determined automatically from the smoothed second derivative of the Ca²⁺ signal using MATLAB's Signal Processing Toolbox. The amplitude of Ca²⁺ transients (amplitude) was determined by numerically integrating the area underneath each Ca²⁺ peak relative to the baseline. Irregularity in both timing and amplitude was defined in the same manner as described above for EB contraction rates.

CREBBP mutations in relapsed acute lymphoblastic leukaemia

Charles G. Mullighan^{1*}, Jinghui Zhang^{2*}, Lawryn H. Kasper³, Stephanie Lerach³, Debbie Payne-Turner¹, Letha A. Phillips¹, Sue L. Heatley¹, Linda Holmfeldt¹, J. Racquel Collins-Underwood¹, Jing Ma⁴, Kenneth H. Buetow^{5,6}, Ching-Hon Pui⁷, Sharyn D. Baker⁸, Paul K. Brindle³ & James R. Downing¹

Relapsed acute lymphoblastic leukaemia (ALL) is a leading cause of death due to disease in young people, but the biological determinants of treatment failure remain poorly understood. Recent genome-wide profiling of structural DNA alterations in ALL have identified multiple submicroscopic somatic mutations targeting key cellular pathways^{1,2}, and have demonstrated substantial evolution in genetic alterations from diagnosis to relapse³. However, DNA sequence mutations in ALL have not been analysed in detail. To identify novel mutations in relapsed ALL, we resequenced 300 genes in matched diagnosis and relapse samples from 23 patients with ALL. This identified 52 somatic non-synonymous mutations in 32 genes, many of which were novel, including the transcriptional coactivators *CREBBP* and *NCOR1*, the transcription factors *ERG*, *SPI1*, *TCF4* and *TCF7L2*, components of the Ras signalling pathway, histone genes, genes involved in histone modification (*CREBBP* and *CTCF*), and genes previously shown^{1,2} to be targets of recurring DNA copy number alteration in ALL. Analysis of an extended cohort of 71 diagnosis–relapse cases and 270 acute leukaemia cases that did not relapse found that 18.3% of relapse cases had sequence or deletion mutations of *CREBBP*, which encodes the transcriptional coactivator and histone acetyltransferase CREB-binding protein (*CREBBP*, also known as *CBP*)⁴. The mutations were either present at diagnosis or acquired at relapse, and resulted in truncated alleles or deleterious substitutions in conserved residues of the histone acetyltransferase domain. Functionally, the mutations impaired histone acetylation and transcriptional regulation of *CREBBP* targets, including glucocorticoid responsive genes. Several mutations acquired at relapse were detected in subclones at diagnosis, suggesting that the mutations may confer resistance to therapy. These results extend the landscape of genetic alterations in leukaemia, and identify mutations targeting transcriptional and epigenetic regulation as a mechanism of resistance in ALL.

Acute lymphoblastic leukaemia (ALL) is the commonest childhood malignancy⁵, and is a leading cause of cancer-related death in young people. Several structural chromosomal alterations in ALL, including rearrangement of *MLL* and the Philadelphia chromosome⁶ are associated with a high risk of treatment failure and relapse. However, many ALL cases that fail therapy lack these alterations, and the biological basis of treatment failure in these cases is poorly understood. Genome-wide profiling of ALL has identified multiple recurring submicroscopic genetic alterations targeting lymphoid development, cell cycle regulation, tumour suppression and apoptosis^{1,2}, and has identified genetic alterations that predict a high risk of relapse, including deletion of *IKZF1* (encoding the transcription factor IKAROS)^{7,8}. Moreover, profiling of structural DNA alterations in matched diagnosis and relapse

samples has demonstrated that in the majority of cases there are substantial differences in the complement of genetic lesions between diagnosis and relapse, although the predominant clones at both stages of disease share a common ancestral origin^{3,9}. The predominant clones at relapse are commonly present at low levels at diagnosis, suggesting that specific genetic alterations may confer resistance to therapy. Frequently acquired lesions at relapse include deletions of *CDKN2A/B*, *ETV6* and *IKZF1*^{3,9}. However, a detailed analysis of sequence variation in ALL has not been performed.

To identify novel sequence mutations in relapsed ALL, we resequenced 300 genes in matched diagnosis–relapse samples from 23 children with B-cell progenitor ALL. Cases studied included B-cell progenitor ALL with high hyperdiploidy ($N = 3$), *TCF3-PBX1* ($N = 1$), *ETV6-RUNX1* ($N = 3$), rearrangement of *MLL* ($N = 3$), *BCR-ABL1* ($N = 3$), and low hyperdiploid, pseudodiploid, or miscellaneous karyotypes ($N = 10$) (Supplementary Table 1). We sequenced genes known to be mutated in leukaemia and cancer, and those that encode pathway components targeted by recurring copy number alteration in ALL, tumour suppressors, cell cycle regulators, tyrosine kinases, and genes encoding DNA repair proteins (Supplementary Table 2). We identified 52 somatic (non-inherited) protein coding variants in 32 genes in 20 cases (mean 2.5 variants per case, range 0–5) (Supplementary Table 3). Somatic mutations were identified in genes previously known to be mutated in haematopoietic malignancies and ALL, including the ETS-family transcription factor gene *ETV6* ($N = 1$); the Janus kinase gene *JAK1* ($N = 1$); the Ras pathway genes *NRAS* ($N = 5$; Supplementary Table 4 and Supplementary Fig. 1), *KRAS* ($N = 2$), *NF1* ($N = 3$) and *PTPN11* ($N = 2$); the B-lymphoid transcription factor gene *PAX5* ($N = 2$); the U3 ubiquitin ligase gene *FBXW7* ($N = 1$); the histone methyltransferase gene *EZH2* ($N = 1$); and the tumour suppressor gene *TP53* ($N = 2$). In addition, we observed patterns of evolution of sequence variations between diagnosis and relapse that recapitulated those observed for DNA copy number alterations in ALL³ (Table 1, Supplementary Results and Supplementary Table 4).

A novel finding was somatic coding mutations in *CREBBP* (or *CBP*, encoding CREB-binding protein) in four of 23 cases sequenced (Supplementary Table 5). *CREBBP* was selected for sequencing by identifying recurring focal deletions involving the gene. *CREBBP* and its paralogue, *EP300* (p300) are transcriptional coactivators with multiple functions in development and haematopoiesis^{4,10,11}. Both are molecular scaffolds that interact with a diverse range of transcription factors, regulate transcription by acetylation of histone and non-histone targets, and may regulate protein turnover by E4 polyubiquitin ligase activity. *CREBBP* and *EP300* are known targets of translocations in acute leukaemia (for example, *MLL-EP300*, *MLL-CREBBP*, *MOZ-CREBBP* and *MOZ-EP300*)¹². Inherited *CREBBP* mutations and

¹Department of Pathology, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ²Department of Computational Biology, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ³Department of Biochemistry, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ⁴The Hartwell Center for Bioinformatics and Biotechnology, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ⁵National Cancer Institute Center for Bioinformatics, National Cancer Institute, Rockville, Maryland 20892, USA. ⁶Laboratory of Population Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. ⁷Department of Oncology, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ⁸Department of Pharmaceutical Sciences, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA.

*These authors contributed equally to this work.

Table 1 | Recurrent mutations in 23 matched diagnosis–relapse ALL cases

Gene	N	Present at diagnosis and relapse	Present at diagnosis only	Present at relapse only	Description/function
<i>ASMTL</i>	3	2		1	Acetylserotonin O-methyltransferase-like
<i>CREBBP</i>	4	3		1	Transcriptional coactivator, histone and non-histone acetyl transferase, ubiquitin ligase
<i>ERG</i>	2			2	ETS family transcription factor
<i>FLT3</i>	2		1	1	Receptor tyrosine kinase, Ras pathway
<i>KRAS</i>	3	2		1	Ras pathway
<i>NF1</i>	2		2		Ras pathway
<i>NRAS</i>	5	2		1	Ras pathway
<i>PAX5</i>	2	1	1		B cell development
<i>PTPN11</i>	2	1		1	Ras pathway
<i>TP53</i>	2			2	Tumour suppressor
<i>TUSC3</i>	2	2			Tumour suppressor candidate

deletions result in the Rubinstein–Taybi syndrome, a developmental disorder characterized by dysmorphology, intellectual impairment, and an increased susceptibility to solid tumours^{13,14}. Homozygous deletion of *Crebbp* or *Ep300* is lethal in mice owing to developmental abnormalities, and *Crebbp*^{+/-} mice show defects in B lymphoid development, and an increased incidence of haematopoietic tumours¹⁵. Although *CREBBP* and *EP300* sequence mutations have been reported in solid tumours^{16,17} and rare *EP300* mutations have been detected in an ALL cell line and myelodysplasia¹⁸, there are no prior reports of *CREBBP* sequence mutations in haematologic malignancies.

To define the frequency of *CREBBP* and *EP300* mutations in acute leukaemia, we sequenced these genes in an additional 318 cases of paediatric leukaemia. These additional cases included matched diagnosis and relapse samples from 48 children with ALL, and diagnosis samples from children with ALL (*N* = 170) and acute myeloid leukaemia (AML,

N = 100)¹⁹ that did not relapse. Single nucleotide polymorphism (SNP) microarray DNA copy number alteration and loss-of-heterozygosity data were available for all cases sequenced^{1,20,21}. We also examined DNA copy number alterations in a further 107 ALL cases that were not sequenced^{1,3,20}. In addition, *CREBBP* was sequenced in 58 ALL and AML cell lines. Remarkably, 13 of 71 (18.3%) of relapsed ALL cases harboured either tumour-acquired (non-inherited) sequence alterations (*N* = 13) or focal deletions (*N* = 2) of *CREBBP* (Supplementary Table 5 and Supplementary Fig. 2). In contrast, *CREBBP* alterations in cases of childhood acute leukaemia that did not relapse were rare, with only one additional *CREBBP* mutation identified in 200 AML and ALL cases sequenced (ALL patient Hyperdiploid-#22, C1408Y, Supplementary Table 5). Furthermore, three of 307 ALL cases with SNP array data had focal *CREBBP* deletions at diagnosis (Supplementary Table 5 and Supplementary Fig. 2). *CREBBP* alterations were not observed in AML.

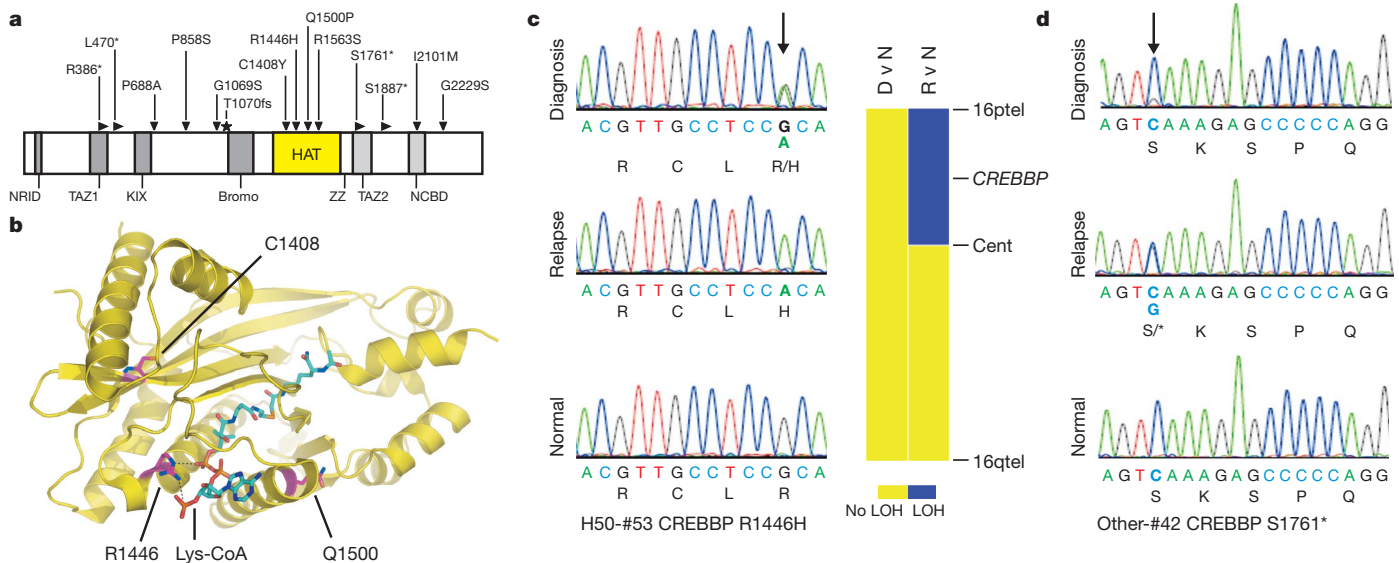


Figure 1 | *CREBBP* sequence mutations in relapsed ALL. **a**, Most variants are missense mutations in *CREBBP* domains involved in histone acetylation or transcription factor recruitment, or result in protein truncation. NRID, nuclear-receptor-interaction domain; TAZ1/2, transcriptional-adaptor zinc-finger 1/2; KIX, KID-binding domain; Bromo, bromodomain; HAT, histone acetyltransferase domain; ZZ, zinc-binding domain near the dystrophin WW domain; NCBD, nuclear-receptor coactivator-binding domain. **b**, The locations of *CREBBP* HAT mutations are shown using the crystal structure of the EP300 HAT domain complexed with its bisubstrate inhibitor, Lys-CoA (blue)²². *CREBBP* R1446 (equivalent to EP300 R1410) contacts phosphates of the CoA moiety of the inhibitor (salt bridges are shown as dashed lines), and the R1446H mutation is predicted to disrupt substrate binding. Q1500P (EP300 Q1464) is predicted to disrupt the α 4 helix, which stabilizes the substrate binding loop L1 between the β 5 strand and the α 4 helix. C1408H (EP300 C1372) is predicted to disrupt the hydrophobic core that involves both the amino and carboxy termini of the HAT domain. R1563 cannot be shown because this residue lies in a proteolytically sensitive autoacetylation loop that

was deleted in order to generate the crystal structure²². **c**, Sequence traces (labelled with nucleotide, and corresponding amino acid residues R, C, L and R/H) showing duplication of the R1446H mutation at relapse for patient H50-#53. This mutation was heterozygous at diagnosis and absent in the matched normal sample. There is copy-neutral loss of heterozygosity (LOH) of chromosome 16p at relapse but not at diagnosis. D v N is the LOH analysis of diagnosis sample compared to matched normal sample; R v N is the relapse compared to matched normal sample. Cent, centromere; 16ptel, chromosome 16p telomere; 16qtel, chromosome 16q telomere. **d**, Sequence traces (labelled with nucleotide, and corresponding amino acid residues S, K, P and Q). In patient Other-#42, *CREBBP* mutations are present in subclones at diagnosis, and emerge in the predominant clone at relapse. The S1761* mutation is heterozygous in the relapse sample, absent in the matched normal sample, and appears as a minor peak in the diagnosis sample. Presence of this mutation in a subpopulation of cells at diagnosis was confirmed by PCR, cloning and bidirectional sequencing of multiple colonies of the diagnosis sample (data not shown).

Only one *EP300* mutation was detected in the 71 cases that relapsed, a missense mutation (P925L) acquired at relapse. The identified *CREBBP* mutations resulted in amino acid substitutions, most commonly in the histone acetyltransferase (HAT) domain, or truncating frameshift or nonsense changes (Fig. 1a). The HAT mutations involved highly conserved residues (Supplementary Fig. 3). Modelling of the predicted effects of the *CREBBP* HAT mutations using the crystal structure of the highly homologous *EP300* HAT domain²² demonstrated that the mutations are likely to disrupt the structure of the domain or its interaction with substrates (Fig. 1b).

Notably, *CREBBP* mutations present at diagnosis were retained or duplicated at relapse. Three cases had biallelic *CREBBP* mutations, either compound heterozygosity for different mutations, or homozygosity for a single mutation (Q1500P). Q1500P, but none of the other mutations identified in ALL, is also observed in Rubinstein–Taybi syndrome¹³. Homozygosity for this mutation was accompanied by DNA copy-neutral loss-of-heterozygosity (acquired uniparental disomy) of chromosome

16p in both diagnosis and relapse samples, but not the germline sample in this case. This was identified on analysis of SNP microarray genotype data for this patient, indicating that one copy of 16p containing the wild-type *CREBBP* allele had been deleted, while the remaining copy of 16p harbouring the Q1500P allele had been duplicated. One case (patient Hyperdiploid-#53) had a heterozygous HAT domain mutation (R1446H) at diagnosis that was homozygous at relapse with accompanying copy-neutral loss-of-heterozygosity of 16p, again indicating duplication of the mutated allele in the predominant relapse clone (Fig. 1c). Furthermore, three mutations detected at relapse (T1070fs, S1761* and I2101M) were detected at low levels in the diagnosis sample (Fig. 1d).

Together, the high frequency of *CREBBP* mutations in relapsed ALL, the persistence, reduplication or emergence of mutations from diagnosis to relapse, and the location of the mutations in key *CREBBP* functional domains suggest that these alterations impair *CREBBP* function and influence treatment responsiveness. *CREBBP* is expressed in

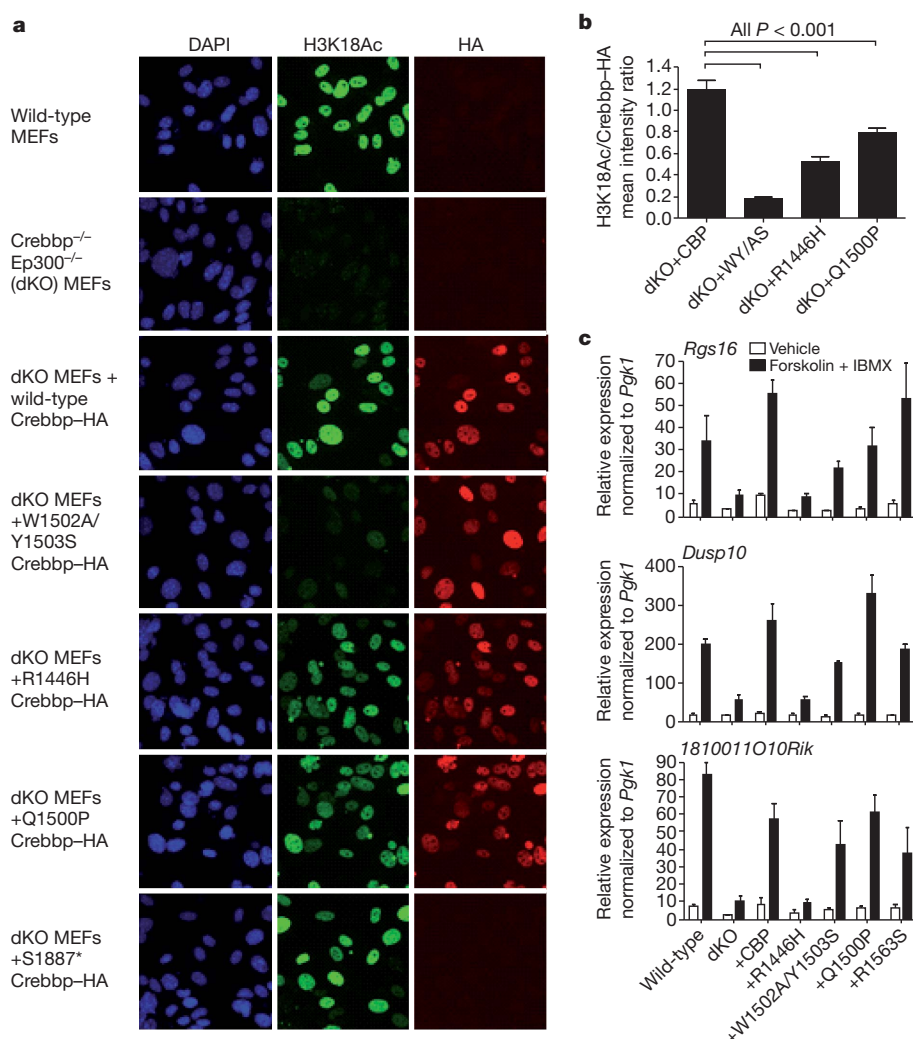


Figure 2 | CREBBP mutations impair histone acetylation and multiple gene expression programs. **a**, Immunofluorescence to detect histone H3 lysine 18 acetylation (H3K18Ac) in nontransduced wild-type MEFs, *Crebbp*^{Aflx/Aflx}; *Ep300*^{Aflx/Aflx} (dKO) MEFs and dKO MEFs transduced with retrovirus expressing wild type (+*Crebbp*) and mutant haemagglutinin (HA)-tagged *Crebbp*. Four independent experiments with separate controls were performed, one of which is shown. S1887* truncates the protein before the C-terminal HA-tag, but is predicted to retain the HAT domain. **b**, Quantification of H3K18Ac mean signal intensity per nucleus relative to the HA-tagged *Crebbp* retrovirus mean signal intensity. Values are mean \pm s.e.m.; 40–61 nuclei quantified per retrovirus. Only nuclei that have an HA-tag

(*Crebbp*-HA) signal greater than 2.5-fold above background were included. Data are expressed as the ratio of the mean H3K18Ac signal intensity for each nucleus to the mean HA signal intensity for the same nuclei. The P value shown is from a Tukey post-test of one-way analysis of variance (ANOVA); mean \pm standard error of the mean. **c**, Quantitative reverse transcription (RT)-PCR gene expression data from dKO MEFs transduced with wild-type or mutant CBP, and treated with ethanol (vehicle) or with 10 μ M forskolin + 100 μ M IBMX for 90 min. W1502A/Y1503S (WY/AS) is a previously described mutation that inactivates HAT enzymatic activity³⁰. Gene expression was normalized to expression of *Pgk1*. $N = 3$ –7.

leukaemic cells and normal B-cell progenitors (Supplementary Fig. 4), and the mutant *CREBBP* alleles are expressed in ALL cell lines harbouring mutations (Supplementary Table 6 and Supplementary Fig. 5).

We investigated the functional effect of the mutations in mice. We examined the histone acetylation and transcriptional activities of wild-type and mutant *Crebbp* alleles expressed in *Crebbp^{Aflx/Aflx};Ep300^{Aflx/Aflx}* (CBP/p300 Cre-deleted double knockout, or dKO) primary murine embryonic fibroblasts (MEFs)^{23–25} (Fig. 2 and Supplementary Figs 6–10). The CREBBP HAT mutations resulted in diminished acetylation of histone H3 lysine 18 (H3K18), a known acetylation substrate of CREBBP²⁵ (Fig. 2a and b). Notably, the acetylation was attenuated but not blocked, and differed between HAT mutants (Fig. 2a and b). The Rubinstein–Taybi syndrome Q1500P mutation resulted in only a modest reduction in acetylation despite being predicted to disrupt a key α helix in the HAT domain (Fig. 1a). This suggests that HAT mutants have additional deleterious effects beyond impairing histone acetylation, that H3K18 is not the critical substrate, or that attenuated HAT activity is sufficient for the phenotype. To investigate these possibilities, we examined the effects of the CREBBP mutations on the expression of *Crebbp* target genes and cell proliferation in dKO MEFs transduced with retrovirus expressing wild-type and mutant *Crebbp* complementary DNAs (Fig. 2c and Supplementary Figs 6–10). We tested mutations in the HAT domain, as well as S1761*, S1887* and I2101M, which are predicted to affect the nuclear coactivator binding domain (Fig. 1a). Interestingly, the HAT mutants result in reduced expression of cyclic-AMP-responsive CREB target genes (Fig. 2c and Supplementary Figs 6–8). Furthermore, many of the mutants tested also impaired cell proliferation (Supplementary Fig. 11). Truncating mutations within the carboxy-terminal region of the HAT domain affected the expression of double-stranded RNA and glucocorticoid-receptor-responsive genes (Supplementary Fig. 8j–k and Supplementary Figs 9–10), consistent with the role of the nuclear coactivator binding domain in these pathways. The magnitude of reduced expression depended on the target gene, the specific mutation and the pathway, but the finding that CREBBP mutations result in impaired expression of glucocorticoid-receptor-responsive genes is particularly notable. The glucocorticoid dexamethasone is a mainstay of therapy for ALL, and poor responsiveness to steroid therapy is strongly associated with poor treatment outcome²⁶. Accordingly, we examined responsiveness to dexamethasone and the class I/II HDAC inhibitor vorinostat in a panel of T-lineage ALL cell lines with wild-type or mutant *CREBBP* alleles. This demonstrated dexamethasone resistance in the majority of cell lines, but sensitivity to vorinostat at clinically useful concentrations (half-maximum inhibitory concentration IC₅₀ below 1 μ M) in the majority of cell lines tested (Supplementary Table 7 and Supplementary Fig. 12). This suggests that multiple mechanisms are likely to influence glucocorticoid responsiveness, but that HDAC inhibitor therapy may be useful in steroid-resistant ALL.

These findings show that detailed analysis of sequence alterations in relapsed ALL can identify novel genetic alterations that are likely to be involved in the pathogenesis of treatment failure. Approximately 10% of the genes sequenced harboured somatic sequence mutations with multiple mutations involving transcriptional regulators and coactivators, many of which are targeted by other structural genetic alterations in ALL, including deletions and translocations (for example, *PAX5*, *ERG*, *ETV6*, *CREBBP* and *EP300*). These results, together with those of Pasqualucci *et al.* identifying a high frequency of *CREBBP* and *EP300* mutations in diffuse large B-cell and follicular lymphoma²⁷, also identify *CREBBP* and *EP300* as new targets of recurring mutation in a range of lymphoid malignancies. The observation that the *CREBBP* mutations impair regulation of glucocorticoid-responsive genes, and that the mutations are selected for at relapse, suggests that these alterations may influence response to therapy and the likelihood of relapse. These results are also of clinical relevance, because they suggest that therapeutic approaches directed at modulating protein acetylation²⁸

may be useful in high-risk ALL, particularly as HDAC inhibitors may induce apoptosis in glucocorticoid-resistant leukaemic cells²⁹. Finally, these results indicate that comprehensive evaluation of sequence alterations and epigenetic modifications in relapsed ALL is likely to yield further biological insights and potential therapeutic approaches for this disease.

METHODS SUMMARY

Sequencing of 300 genes was performed by polymerase chain reaction (PCR) and capillary resequencing of whole genome amplified DNA extracted from leukaemic cells obtained at diagnosis and relapse from 23 B-cell progenitor ALL cases. *CREBBP* and *EP300* mutation recurrence testing was performed in an additional 48 diagnosis–relapse B-cell and T-cell ALL samples, and *CREBBP* was also sequenced in 270 ALL and AML samples that did not relapse. All putative mutations were validated as somatic by sequencing of corresponding remission DNA samples, and by sequencing of unamplified tumour and matched normal DNA. All cases had DNA copy number alteration data from Affymetrix 500K or SNP 6.0 microarrays.

In vitro functional assays of CREBBP mutants. H3K18 acetylation, gene expression and cell proliferation were assayed in *Crebbp^{Aflx/Aflx};Ep300^{Aflx/Aflx}* dKO MEFs^{23–25} transduced with retroviral supernatants expressing wild-type and mutant *Crebbp* alleles. Dexamethasone and vorinostat drug response assays and all methods are described in full in the online-only Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 June; accepted 1 December 2010.

- Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758–764 (2007).
- Kuiper, R. P. *et al.* High-resolution genomic profiling of childhood ALL reveals novel recurrent genetic lesions affecting pathways involved in lymphocyte differentiation and cell cycle progression. *Leukemia* **21**, 1258–1266 (2007).
- Mullighan, C. G. *et al.* Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**, 1377–1380 (2008).
- Goodman, R. H. & Smolik, S. CBP/p300 in cell growth, transformation, and development. *Genes Dev.* **14**, 1553–1577 (2000).
- Pui, C. H., Robison, L. L. & Look, A. T. Acute lymphoblastic leukaemia. *Lancet* **371**, 1030–1043 (2008).
- Harrison, C. J. Cytogenetics of paediatric and adolescent acute lymphoblastic leukaemia. *Br. J. Haematol.* **144**, 147–156 (2009).
- Mullighan, C. G. *et al.* Deletion of *IKZF1* and prognosis in Acute Lymphoblastic Leukemia. *N. Engl. J. Med.* **360**, 470–480 (2009).
- Kuiper, R. P. *et al.* *IKZF1* deletions predict relapse in uniformly treated pediatric precursor B-ALL. *Leukemia* **24**, 1258–1264 (2010).
- Yang, J. J. *et al.* Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukemia. *Blood* **112**, 4178–4183 (2008).
- Vo, N. & Goodman, R. H. CREB-binding protein and p300 in transcriptional regulation. *J. Biol. Chem.* **276**, 13505–13508 (2001).
- Blobel, G. A. CREB-binding protein and p300: molecular integrators of hematopoietic transcription. *Blood* **95**, 745–755 (2000).
- Yang, X. J. The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. *Nucleic Acids Res.* **32**, 959–976 (2004).
- Schorr, E. K. *et al.* Genotype-phenotype correlations in Rubinstein-Taybi syndrome. *Am. J. Med. Genet. A* **146A**, 2512–2519 (2008).
- Miller, R. W. & Rubinstein, J. H. Tumors in Rubinstein-Taybi syndrome. *Am. J. Med. Genet.* **56**, 112–115 (1995).
- Kung, A. L. *et al.* Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes Dev.* **14**, 272–277 (2000).
- Iyer, N. G., Ozdag, H. & Caldas, C. p300/CBP and cancer. *Oncogene* **23**, 4225–4231 (2004).
- Kishimoto, M. *et al.* Mutations and deletions of the CBP gene in human lung cancer. *Clin. Cancer Res.* **11**, 512–519 (2005).
- Shigeno, K. *et al.* Disease-related potential of mutations in transcriptional cofactors CREB-binding protein and p300 in leukemias. *Cancer Lett.* **213**, 11–20 (2004).
- Radtke, I. *et al.* Genomic analysis reveals few genetic alterations in pediatric acute myeloid leukemia. *Proc. Natl Acad. Sci. USA* **106**, 12944–12949 (2009).
- Mullighan, C. G. *et al.* *BCR-ABL1* lymphoblastic leukaemia is characterized by the deletion of *Ikaros*. *Nature* **453**, 110–114 (2008).
- Mullighan, C. G. *et al.* Rearrangement of *CRLF2* in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. *Nature Genet.* **41**, 1243–1246 (2009).
- Liu, X. *et al.* The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature* **451**, 846–850 (2008).
- Kang-Decker, N. *et al.* Loss of CBP causes T cell lymphomagenesis in synergy with p27Kip1 insufficiency. *Cancer Cell* **5**, 177–189 (2004).

24. Kasper, L. H. *et al.* Conditional knockout mice reveal distinct functions for the global transcriptional coactivators CBP and p300 in T-cell development. *Mol. Cell. Biol.* **26**, 789–809 (2006).
25. Kasper, L. H. *et al.* CBP/p300 double null cells reveal effect of coactivator level and diversity on CREB transactivation. *EMBO J.* **29**, 3660–3672 (2010).
26. Dordelmann, M. *et al.* Prednisone response is the strongest predictor of treatment outcome in infant acute lymphoblastic leukemia. *Blood* **94**, 1209–1217 (1999).
27. Pasqualucci, L. *et al.* Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* doi:10.1038/nature09730 (this issue).
28. Bolden, J. E., Peart, M. J. & Johnstone, R. W. Anticancer activities of histone deacetylase inhibitors. *Nature Rev. Drug Discov.* **5**, 769–784 (2006).
29. Tsapis, M. *et al.* HDAC inhibitors induce apoptosis in glucocorticoid-resistant acute lymphatic leukemia cells despite a switch from the extrinsic to the intrinsic death pathway. *Int. J. Biochem. Cell Biol.* **39**, 1500–1509 (2007).
30. Bordoli, L. *et al.* Functional analysis of the p300 acetyltransferase domain: the PHD finger of p300 but not of CBP is dispensable for enzymatic activity. *Nucleic Acids Res.* **29**, 4462–4471 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Jeevan, S. Orwick and A. Gibson for technical assistance, B. Schulman for assistance with structural modelling, and B. Woolf and

J. Hartigan of Beckman Coulter Genomics for assistance with sequencing. We thank the Tissue Resources Facility of St Jude Children's Research Hospital for providing samples, and the following St Jude core facilities: Vector Development and Production, Flow Cytometry and Cell Sorting, Cell and Tissue Imaging, the Animal Resource Center, and the DNA sequencing and Macromolecular Synthesis laboratories of the Hartwell Center for Bioinformatics and Biotechnology. This study was supported by ALSAC of St Jude and Cancer Center support grant P30 CA021765, and grant number DE018183 (P.K.B.). C.G.M. is a Pew Scholar in the Biomedical Sciences.

Author Contributions C.G.M., P.K.B. and J.R.D. designed the study. S.L.H., L.H., C.G.M., L.A.P. and D.P.-T. performed PCR and sequencing. J.Z. and K.H.B. analysed sequence data. L.H.K. and S.L. performed *in vitro* assays of the functional activity of Crebbp mutants. J.M. analysed genomic data. S.L.H. and J.R.C.-U. performed cell line assays. S.D.B. designed and performed leukaemia cell line drug responsiveness assays. C.-H.P. provided samples and clinical data. C.G.M. wrote the manuscript. All authors reviewed the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.G.M. (charles.mullighan@stjude.org).

METHODS

Patients and samples. Seventy-one children with matched diagnosis and remission ALL samples were studied, including 61 cases previously studied by SNP microarray analysis³. All samples had at least 80% blasts by immunophenotypic and/or morphologic analysis, or were flow-sorted to at least 90% purity before DNA extraction. Sequencing of 300 genes was performed for 23 B-cell progenitor ALL cases. *CREBBP* and *EP300* mutation recurrence testing was performed in an additional 48 diagnosis-relapse B-cell and T-cell ALL samples. Samples obtained at diagnosis from 270 children with AML and ALL who did not experience relapse were also sequenced for *CREBBP* variants. This cohort comprised AML cases with translocation (8;21) [*RUNX1-RUNX1T1*] (*N* = 9), inversion/translocation (16;16) [*CBFB-SMMHC*] (*N* = 16), translocation (15;17) or related rearrangements involving *RARA* (*N* = 7), M7 morphology (*N* = 9), rearrangement of *MLL* (*N* = 15), normal or miscellaneous karyotype (*N* = 44); B-cell progenitor ALL cases with high hyperdiploidy (*N* = 13), *ETV6-RUNX1* (*N* = 13), *TCF3-PBX1* (*N* = 13), rearrangement of *MLL* (*N* = 12), *BCR-ABL1* (*N* = 15), hypodiploidy with 44–45 chromosomes (*N* = 10), normal, pseudodiploid or miscellaneous karyotype (*N* = 12), and T-cell lineage ALL (*N* = 82). The study was approved by the St Jude Children's Research Hospital Institutional Review Board.

CREBBP was also sequenced in 58 AML and ALL cell lines. Cell lines were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen, the American Type Culture Collection, from local institutional repositories, or were gifts from O. Heidenreich (SKNO-1) and D. Campana (OP1). Cells were cultured in accordance with previously published recommendations³¹. The paediatric *BCR-ABL1* B-cell precursor ALL cell line OP1 (ref. 32) was cultured in RPMI-1640 containing 100 units per millilitre penicillin, 100 µg ml⁻¹ streptomycin, 2 mM glutamine and 10% fetal bovine serum. DNA was extracted from 5 × 10⁶ cells obtained during log phase growth after washing in PBS using the QIAamp DNA blood mini kit (Qiagen).

The ALL cell lines were as follows: 380 (*MYC-IGH* and *BCL2-IGH* B-cell precursor), 697 (*TCF3-PBX1*), ALL-SIL (T-cell ALL with *HOX11-TCRD* and *NUP214-ABL1*), AT1 (*ETV6-RUNX1*), B1 (*MLL-AF4*), BV173 (chronic myelogenous leukaemia in lymphoid blast crisis), CCRF-CEM (T-cell ALL with *TAL-SIL*), CTV-1 (T-cell ALL with *TAL1-TRB*), DND41 (T-cell ALL with *HOX11L2/TLX3-BCL11B* alteration), HPB-ALL (T-cell ALL with *HOX11L2/TLX3-BCL11B*), HSB-2 (T-cell ALL), Jurkat (T-cell ALL), Kasumi-2 (*TCF3-PBX1*), KARPAS 45 (T-cell ALL with *MLL-MLLT7*), KE-37 (*MYC-TRAD*), KOPT-K1 (T-cell ALL with *LMO2/TTG2-TRD*), LOUCY (T-cell ALL), MHH-CALL-2 (hyperdiploid B-cell precursor ALL), MHH-CALL-3 (*TCF3-PBX1*), MHH-CALL-4 (*IGH@-CRLF2*), MOLT3 (T-cell ALL), MOLT 13 (T-cell ALL), MOLT 15 (T-cell ALL; may be identical to CTV-1), MKB-1 (T-cell ALL line; may be subline of CCRF-CEM), MOLT4 (T-cell ALL), MUTZ5 (*IGH@-CRLF2*), NALM-6 (B-cell precursor ALL), OP1 (*BCR-ABL1*), P12/Ichikawa (T-cell ALL), PEER (T-cell ALL with *NKX2-5-BCL11B* and *NUP214-ABL1*), PF-382 (T-cell ALL), Reh (*ETV6-RUNX1*), REX (T-cell ALL), RPMI 8402 (T-cell ALL with *LMO1/TTG1-TRAD*, *SIL-TAL1/SCL*), RS4;11 (*MLL-AF4*), SD1 (*BCR-ABL1*), SKW3 (T-cell ALL, may represent KE-37), SUP-B15 (*BCR-ABL1*), SUP-T1 (T-cell ALL with *TRB-NOTCH1*), SUP-T11 (T-cell ALL with *TRA/D-TCL1*), TALL-1 (T-cell ALL), TOM-1 (*BCR-ABL1*), U-937 (*PICALM-AF10*), UOCB1 (*TCF3-HLF*) and YT (natural killer cell leukaemia).

The AML cell lines were as follows: CMK (FAB M7), HL-60 (FAB M2), K-562 (chronic myelogenous leukaemia in myeloid blast crisis), Kasumi-1 (*RUNX1-RUNX1T1*), KG-1 (myelocytic leukaemia), ME-1 (*CBFB-MYH11*), ML-2 (*MLL-AF6*), M-07e (FAB M7), Mono Mac 6 (*MLL-AF9*), MV4-11 (*MLL-AF4*), NB4 (*PML-RARA*), NOMO-1 (*MLL-AF9*), PL-21 (FAB M3), SKNO-1 (*RUNX1-RUNX1T1*) and THP-1 (FAB M5).

Genomic sequencing. DNA from diagnosis, relapse and remission samples was amplified by Phi29 polymerase (RepliG, Qiagen). Resequencing of all coding exons was performed by PCR and capillary sequencing. Initial sequencing was performed by Beckman Coulter Genomics (formerly Agencourt Bioscience and Cogenics Danvers). Mutations were validated by sequencing whole genome amplified remission DNA for variants not present in the SNP database, dbSNP³³. *CREBBP* and *EP300* variants not present in matched germline samples were validated by repeat PCR and sequencing of unamplified DNA. PCR primer sequences, PCR reaction conditions, thermal cycling parameters and sequence traces are available upon request. Expression of *CREBBP* variants was confirmed by reverse transcription using Superscript III (Life Technologies) and PCR using either Phusion HF DNA polymerase (New England Biolabs) or BD Advantage DNA polymerase (Clontech). To detect low levels of relapse-acquired mutations in diagnosis samples, (RT-)PCR products were cloned into pGEM-T-Easy (Promega) and at least 24 colonies were bidirectionally sequenced.

Sequence analysis. Base calls and quality scores were determined using the program PHRED^{34,35}. Sequence variations including substitutions and insertion/

deletions (indel) were analysed using the SNPdetector³⁶ and the IndelDetector³⁷ software. A usable read was required to have at least one 30-base pair (bp) window in which 90% of the bases have a PHRED quality score of at least 30. Poor-quality reads were filtered before variation detection. The minimum thresholds of secondary-to-primary peak ratio for substitution and indel detection were set to be 20% and 10%, respectively. All sequence variations were annotated using a previously developed variation annotation pipeline³⁸. Any variation that did not match a known polymorphism (defined as a dbSNP record that does not belong to the OMIM SNP nor COSMIC somatic variation database^{33,39}) and resulted in a non-silent amino acid change was considered a putative mutation. Sequence variants were visualized using Consed⁴⁰.

Western blotting. Western blotting of *CREBBP*/*Crebbp* of either whole-cell lysates or nuclear extracts⁴¹ of leukaemia cell lines and MEFs was performed using anti-HA and anti-*CREBBP* (A22, Santa Cruz Biotechnology) antibodies.

Structural modelling of *CREBBP* and *EP300* HAT domain mutations. The HAT domain amino acid sequences of *CREBBP* and *EP300* are highly homologous. The mutations in the HAT domain in *EP300*, and the homologous *EP300* mutations corresponding to the *CREBBP* HAT mutations, were visualized using the solved crystal structure of a semi-synthetic heterodimeric *EP300* HAT domain in complex with a bi-substrate inhibitor, Lys-CoA²² (<http://dx.doi.org/10.2210/pdb3biy/pdb>)⁴². Visualization was performed using Pymol⁴³.

In vitro analyses of the effects of *CREBBP* mutations on target gene expression, histone acetylation, and cell proliferation MEF isolation and culture were performed as previously described²⁵. MEF treatments used were for 90 min with 10 µM forskolin + 100 µM IBMX or ethanol vehicle, 4 h with 1 µM dexamethasone or ethanol, or 1 h treatment with 100 µg ml⁻¹ Poly I:C or PBS followed by a wash, medium change and incubation for 3 h before harvesting of cells into TRIzol (Life Technologies). Generation of mouse *Crebbp* (CBP)-HA retroviral constructs and retroviral transduction were performed as previously described²⁵. Residue position numbering in the CBP-HA constructs is based on the conserved residue in human *CREBBP*. For all retroviral experiments except proliferation assays, retroviral transductions were 70% or higher. Quantitative RT-PCR, western blot and immunofluorescence protocols were described previously²⁵. H3K18Ac (ab1191) (Abcam), HA-11 monoclonal antibody against the HA epitope (Boehringer Mannheim), *CREBBP* (A-22) and *EP300* (N-20) (Santa Cruz Biotechnology) were used for immunofluorescence. CBP/p300 antiserum (2574) used for immunofluorescence was generated against GST-p300 1-328, but detects both CBP and p300 similarly by immunofluorescence. For proliferation assays, *Crebbp*^{flax/flax}; *Ep300*^{flax/flax}; YFP MEFs were infected with CBP retrovirus, then endogenous *CBP* and *p300* were deleted with adenovirus-Cre after 2 days (day 1 being the day following overnight adenovirus-Cre treatment). Yellow fluorescent protein (YFP) expression was activated in deleted cells. Proliferation assays were commenced with one million YFP⁺ MEFs on day 1 and were passaged every 2–3 days. The total number of YFP⁺ cells on day 11 was calculated from the total cell number and from YFP⁺ percentage as determined by flow cytometry.

Gene expression profiling of MEFs. Total RNA was extracted from untransduced dKO MEFs and dKO MEFs transduced with wild-type and mutant *Crebbp* alleles treated with vehicle or 1 µM dexamethasone for 4 h. RNA was processed and hybridized to HT MG-430 PM gene expression arrays (Affymetrix) according to the manufacturer's instructions. Probe sets were selected for visualization and validation that meet the criteria of threefold induction by dexamethasone in all four dKO + *Crebbp*-HA + dexamethasone samples and which had a minimal level of expression that is above twice the background (≥ 3.1) for the same four samples.

Dexamethasone and vorinostat drug response assays. The activities of dexamethasone (APP Pharmaceuticals) and vorinostat (Selleck Chemicals) were evaluated against nine T-cell ALL cell lines (Jurkat, HPB-ALL, SUPT1, MKB1, REX, P12/Ichikawa, TALL-1, LOUCY and PEER) using an MTT ((3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) cell viability assay. Cells were treated with the indicated drug concentration for 72 h. Data were analysed as the mean percentage of dimethyl sulphoxide (DMSO)-treated control cells at each concentration. IC₅₀ values were calculated from the dose-response curves using nonlinear regression analysis as implemented in the computer software program Prism version 5.0 (GraphPad Software). For each drug and cell line, three to four independent experiments were performed with six replicates at each concentration.

1. Drexler, H. G. *The Leukemia-Lymphoma Cell Line Facts Book* 1st edn (Academic Press, 2001).
2. Manabe, A. *et al.* Interleukin-4 induces programmed cell death (apoptosis) in cases of high-risk acute lymphoblastic leukemia. *Blood* **83**, 1731–1737 (1994).
3. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).

34. Ewing, B., Hillier, L., Wendl, M. C. & Green, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175–185 (1998).
35. Ewing, B. & Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**, 186–194 (1998).
36. Zhang, J. *et al.* SNPdetector: a software tool for sensitive and accurate SNP detection. *PLOS Comput. Biol.* **1**, e53 (2005).
37. Zhang, J. *et al.* Systematic analysis of genetic alterations in tumors using Cancer Genome WorkBench (CGWB). *Genome Res.* **17**, 1111–1117 (2007).
38. Zhang, J., Rowe, W. L., Struewing, J. P. & Buetow, K. H. HapScope: a software system for automated and visual analysis of functionally annotated haplotypes. *Nucleic Acids Res.* **30**, 5213–5221 (2002).
39. Bamford, S. *et al.* The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br. J. Cancer* **91**, 355–358 (2004).
40. Gordon, D., Albajian, C. & Green, P. Consed: a graphical tool for sequence finishing. *Genome Res.* **8**, 195–202 (1998).
41. Andrews, N. C. & Faller, D. V. A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic Acids Res.* **19**, 2499 (1991).
42. Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank. *Nature Struct. Biol.* **10**, 980 (2003).
43. DeLano, W. L. The PyMOL Molecular Graphics System. (<http://www.pymol.org>) (2002).

DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair

Yankun Gao^{1*}, Sachin Katyal^{1*}, Youngsoo Lee¹, Jingfeng Zhao¹, Jerold E. Reh², Helen R. Russell¹ & Peter J. McKinnon¹

DNA replication and repair in mammalian cells involves three distinct DNA ligases: ligase I (Lig1), ligase III (Lig3) and ligase IV (Lig4)¹. Lig3 is considered a key ligase during base excision repair because its stability depends upon its nuclear binding partner Xrcc1, a critical factor for this DNA repair pathway^{2,3}. Lig3 is also present in the mitochondria, where its role in mitochondrial DNA (mtDNA) maintenance is independent of Xrcc1 (ref. 4). However, the biological role of Lig3 is unclear as inactivation of murine Lig3 results in early embryonic lethality⁵. Here we report that Lig3 is essential for mtDNA integrity but dispensable for nuclear DNA repair. Inactivation of Lig3 in the mouse nervous system resulted in mtDNA loss leading to profound mitochondrial dysfunction, disruption of cellular homeostasis and incapacitating ataxia. Similarly, inactivation of Lig3 in cardiac muscle resulted in mitochondrial dysfunction and defective heart-pump function leading to heart failure. However, Lig3 inactivation did not result in nuclear DNA repair deficiency, indicating essential DNA repair functions of Xrcc1 can occur in the absence of Lig3. Instead, we found that Lig1 was critical for DNA repair, but acted in a cooperative manner with Lig3. Additionally, Lig3 deficiency did not recapitulate the hallmark features of neural Xrcc1 inactivation such as DNA damage-induced cerebellar interneuron loss⁶, further underscoring functional separation of these DNA repair factors. Therefore, our data reveal that the critical biological role of Lig3 is to maintain mtDNA integrity and not Xrcc1-dependent DNA repair.

Lig3 has been ascribed repair functions during DNA replication, DNA strand break repair through the BER/single-strand break repair pathway, and has been implicated in double-strand break repair if non-homologous end-joining is disabled^{1,7–9}. Given the close association between Xrcc1 and Lig3 and the substantial DNA repair deficiency present when Xrcc1 is inactivated, the link between Lig3 and DNA repair is compelling^{1–3,7,8}.

In the mitochondria, Lig3 functions during replication and repair of mtDNA, but in an Xrcc1-independent manner^{4,10–12}. The highly oxidative environment of the mitochondria suggests that DNA repair will be important for mtDNA maintenance^{11–17}, particularly as mtDNA point mutations or deletions can result in mitochondrial dysfunction^{13,18–20}. Defects in mtDNA maintenance are the bases of multiple human syndromes. For example, distinct mutations of DNA polymerase gamma (PolG), a mitochondrial-specific polymerase that is essential for mtDNA replication, can result in diverse human syndromes involving ataxia, Parkinsonism or diabetes^{20–24}. Consequently, Lig3 is potentially critical for both nuclear and mtDNA integrity.

To determine the biological role of Lig3, we generated *Lig3^{Nes-cre}* mice, in which *Lig3* was conditionally inactivated throughout the nervous system using *Nestin-cre*. Although germline *Lig3* deletion is lethal early during embryonic development⁵, *Lig3^{Nes-cre}* mice were born alive. Although initially indistinguishable from wild-type (WT) littermates, *Lig3^{Nes-cre}* mice became growth retarded and profoundly ataxic by 2 weeks of age (Supplementary Movie 1), and did not survive beyond 20 days of age (Supplementary Fig. 1). Magnetic resonance imaging of

postnatal day 14 (P14) *Lig3^{Nes-cre}* mice revealed a smaller brain compared with controls (Fig. 1a). *Lig3^{Nes-cre}* tissue had markedly reduced Lig3 levels in the brain, but not elsewhere (Fig. 1b), and despite Lig3 instability after Xrcc1 loss^{2,6}, the converse did not occur (Fig. 1b). In contrast to WT and *Xrcc1^{Nes-cre}*, the *Lig3^{Nes-cre}* cerebellum was substantially smaller, showing both reduced proliferation and increased apoptosis of granule neuron progenitors, resulting in widespread loss of these cells (Fig. 1c, d and Supplementary Fig. 2). Other cell types such as oligodendrocytes were also affected in the *Lig3^{Nes-cre}* brain (Supplementary Fig. 3).

Given that Xrcc1 stabilizes Lig3, we determined if Lig3 loss also disrupts cerebellar interneurons, a hallmark phenotype after Xrcc1 inactivation⁶. However, the *Lig3^{Nes-cre}* cerebellum contained a normal complement of interneurons as determined by western blot of cerebellar extracts and tissue immunostaining using anti-Pax2, a marker for this neuronal type (Fig. 1e, f). Therefore, despite their close association, inactivation of Lig3 leads to a strikingly different neural phenotype to Xrcc1 loss.

As *Xrcc1^{Nes-cre}* and *Lig3^{Nes-cre}* are phenotypically distinct, and because Lig3 inactivation can compromise mtDNA¹¹ and mitochondrial Lig3 is required for ES cell viability³¹, we determined if the *Lig3^{Nes-cre}* neurodevelopmental defects were due to disruption of mitochondrial function. Using mitotracker to label mitochondria, we found a pronounced mitochondrial defect in primary *Lig3^{Nes-cre}* cortical astrocytes (Fig. 2a). In both WT and *Xrcc1^{Nes-cre}* astrocytes there was an elaborate distribution of actively streaming mitochondria. In stark contrast, mitochondria in *Lig3^{Nes-cre}* astrocytes were static and showed an abnormal morphology and distribution (Fig. 2a and Supplementary Movie 2).

Identification of mtDNA using PicoGreen showed a normal distribution in WT or *Xrcc1^{Nes-cre}* cells (arrows), but an absence of mtDNA in *Lig3^{Nes-cre}* astrocytes (Fig. 2a). We then assessed mtDNA integrity *in vivo* in the *Lig3^{Nes-cre}* brain and found the amount of recovered mtDNA from this tissue was markedly decreased compared with that from equal amounts of WT tissue or *Lig3^{Nes-cre}* liver (Fig. 2b). This was further confirmed using a PCR assay that amplified specific mtDNA regions, which showed a loss of mtDNA throughout the *Lig3^{Nes-cre}* brain (Fig. 2b). *Lig3^{Nes-cre}* astrocyte cultures also became acidic compared with WT cultures, suggesting lactic acid accumulation because of defective oxidative respiration, and, accordingly, *Lig3^{Nes-cre}* cells were defective in oxygen metabolism (Supplementary Fig. 4).

We then directly assessed mitochondria ultrastructure in the *Lig3^{Nes-cre}* brain using electron microscopy. We found morphological abnormalities of the mitochondria at postnatal times coinciding with the emergence of ataxia. By P7, cerebellar Purkinje neurons showed mitochondria with distorted cristae structure and broad changes in mitochondrial morphology (Fig. 2c).

We also assessed components of the mtDNA-encoded electron transport chain, which are often affected in dysfunctional mitochondria²². By P5, there was a marked reduction of complex III (CO III), and COX IV immunostaining in Purkinje cells (Fig. 2d), and elsewhere throughout the brain including hippocampal structures and the cortices

¹Department of Genetics, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ²Department of Pathology, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA.

*These authors contributed equally to this work.

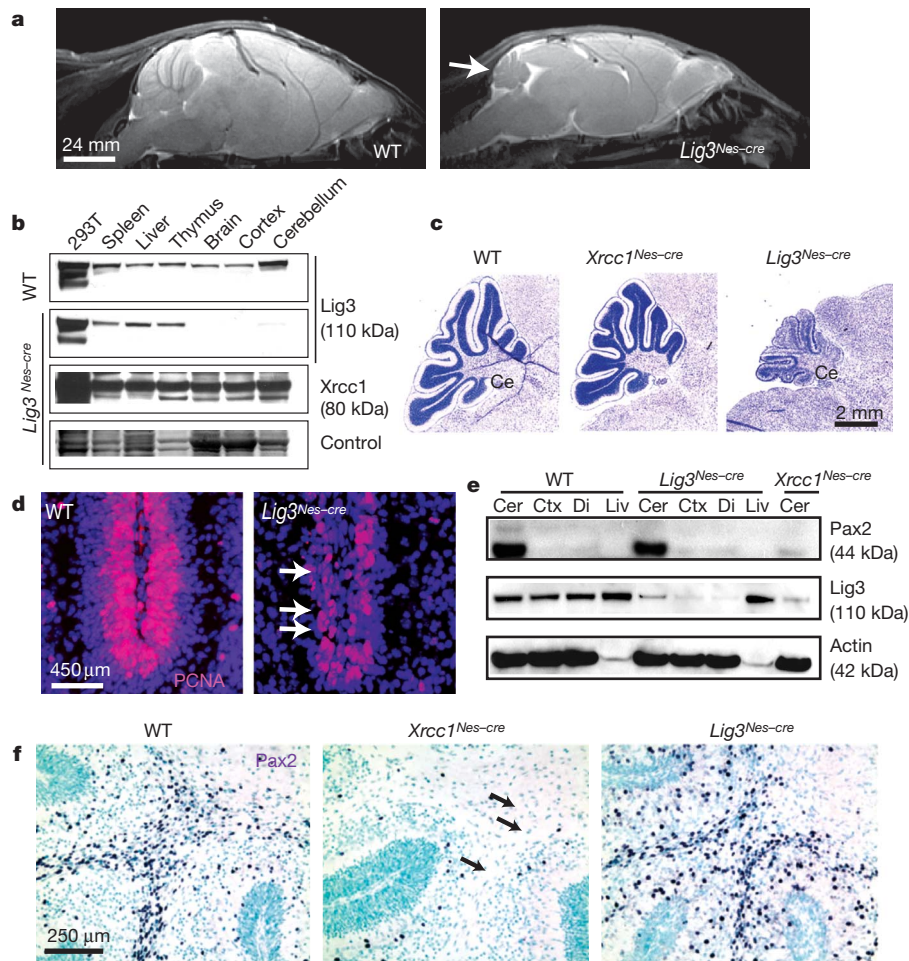


Figure 1 | Lig3 inactivation throughout the nervous system leads to a phenotype different to *Xrcc1* loss. **a**, MRI analysis of the *Lig3^{Nes-cre}* nervous system shows microcephaly and a small cerebellum (arrow) compared with WT. **b**, Lig3 is deleted in various *Lig3^{Nes-cre}* brain regions whereas *Xrcc1* protein levels are unaffected. **c**, Comparative Nissl staining of P14 cerebellum (Ce) from *Lig3^{Nes-cre}*, *Xrcc1^{Nes-cre}* and WT. **d**, PCNA immunostaining shows disrupted neurogenesis in the P6 *Lig3^{Nes-cre}* cerebellum (arrows). **e**, Levels of the interneuron marker Pax2 are similar in control and *Lig3^{Nes-cre}* tissue, but substantially reduced in the *Xrcc1^{Nes-cre}* cerebellum. Cer, cerebellum; Ctx, cortex; Di, diencephalon; Liv, liver. **f**, Interneurons are present in the *Lig3^{Nes-cre}* cerebellum but are absent from *Xrcc1^{Nes-cre}* tissue (arrows) as determined by Pax2 immunostaining.

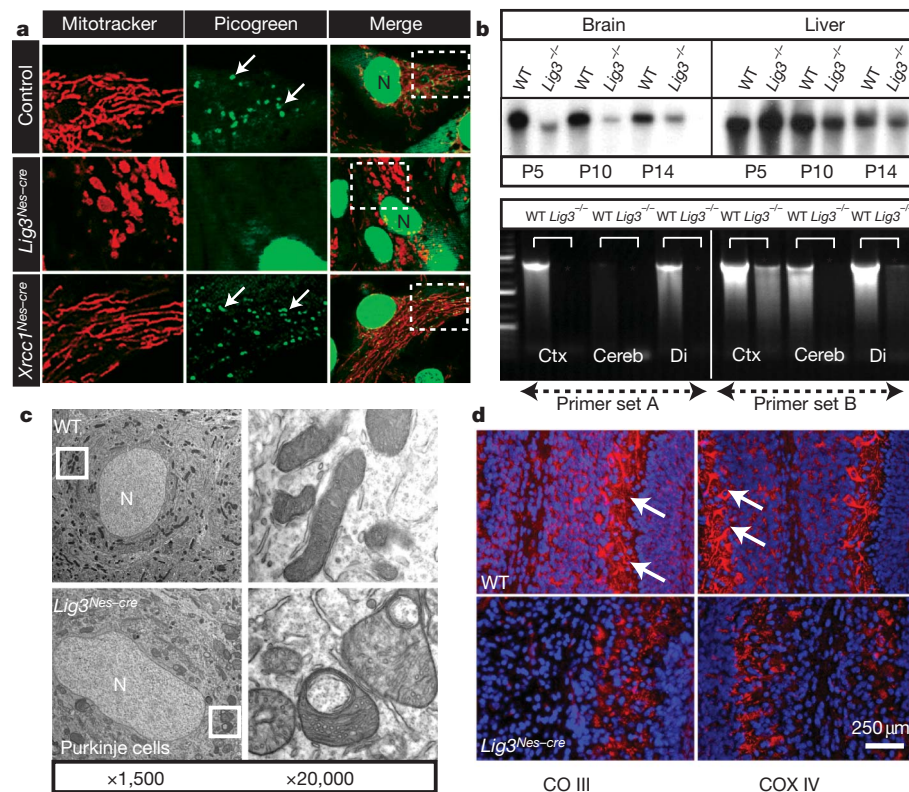


Figure 2 | Mitochondrial function is disrupted in the *Lig3^{Nes-cre}* brain. **a**, Mitotracker labelling reveals defects in mitochondria from *Lig3^{Nes-cre}* but not WT or *Xrcc1^{Nes-cre}* astrocytes. PicoGreen staining shows reduced mtDNA in *Lig3^{Nes-cre}* astrocytes; N, nucleus. Boxed areas show the region expanded in adjacent panels. **b**, Analysis of mtDNA by pulsed-field gel electrophoresis shows an age-dependent quantitative reduction in mtDNA from P5 to P14 *Lig3^{Nes-cre}* brain compared with controls. mtDNA levels in the *Lig3^{Nes-cre}* liver were similar to control tissue. PCR of mtDNA revealed a reduction in mtDNA isolated from *Lig3^{Nes-cre}* brain tissue (asterisks). **c**, Electron microscopy showed alterations in mitochondrial morphology in *Lig3*-deficient Purkinje cells compared with WT tissue. **d**, Staining of electron transport chain components complex III core 2 (CO III) and COX IV *in vivo* is reduced in P7 cerebellar Purkinje cell mitochondria; counterstain is 4',6-diamidino-2-phenylindole (DAPI).

(Supplementary Fig. 5). The biochemical activity of complex I was also reduced in *Lig3^{Nes-cre}* brain (Supplementary Fig. 5). Consistent with increased production of free radicals from dysfunctional mitochondria, increased tyrosine nitrosylation was found in the *Lig3^{Nes-cre}* brains (Supplementary Fig. 6).

We also generated *Lig3^{Ckmm-cre}* mice, in which the muscle creatine kinase promoter drives cre-mediated deletion of *Lig3*, and confirmed diminished *Lig3* levels in cardiac and skeletal muscle, whereas levels in the brain were normal (Fig. 3a). All *Lig3^{Ckmm-cre}* mice died abruptly between 3.5 and 4.5 weeks of age (Fig. 3b). Although we did not find any histological difference between *Lig3* mutant and WT skeletal muscle, *Lig3^{Ckmm-cre}* mice showed heart defects associated with cardiac *Lig3* inactivation. Overall heart size of the mutants was larger than controls, and dilation of the ventricle and atrium and thinning of the ventricle walls was apparent (Supplementary Fig. 7). We also identified disruption in connexin-43, depletion of cardiac troponin-I and defective intercalated disc structure, suggesting that force generation and transmission during cardiac muscle contraction was defective (Fig. 3c and Supplementary Fig. 7). Electron microscopy of mutant heart muscle also revealed disruption of myofibre structure and pronounced defects in mitochondrial morphology (Fig. 3d and Supplementary Fig. 7).

We monitored heart function using transthoracic ultrasound echocardiography (ECG) in *Lig3^{Ckmm-cre}* and WT animals beginning from

3 weeks of age. M-mode echocardiography showed that the diastolic and systolic movement of the left ventricle wall and interventricular septum was decreased in *Lig3^{Ckmm-cre}* mice, which was associated with an ECG reading reflecting an abnormal heart rhythm (Fig. 3e). Fractional shortening and ejection fraction (Fig. 3f), two ways of measuring cardiac contractility, were also decreased significantly in *Lig3^{Ckmm-cre}* mice, indicating heart pump function was severely affected, a phenomenon often leading to heart failure.

Although *Lig3* is linked to DNA repair through interaction with and stabilization by *Xrcc1*^{2,6}, the role for this ligase during nuclear DNA repair is unclear. Therefore, we initially compared nuclear DNA repair in WT, *Lig3^{Nes-cre}* and *Xrcc1^{Nes-cre}* astrocytes after treatment with various genotoxins. We found that DNA repair after hydrogen peroxide and ionizing radiation in quiescent *Lig3^{Nes-cre}* astrocytes was similar to WT cells whereas *Xrcc1^{Nes-cre}* cells showed a marked repair defect (Fig. 4a, b). One exception was ultraviolet damage, where DNA repair in *Lig3^{Nes-cre}* cells was reduced compared with controls; although, even in this situation, *Xrcc1^{Nes-cre}* cells showed a substantially greater DNA repair defect (Supplementary Fig. 8). This may indicate that *Lig3* is selectively required to repair specific types of DNA lesion²⁵.

As *Lig3* loss did not substantially affect DNA repair after hydrogen peroxide or ionizing radiation, we considered if a sub-pathway of BER, termed long-patch-repair⁷, might functionally compensate for

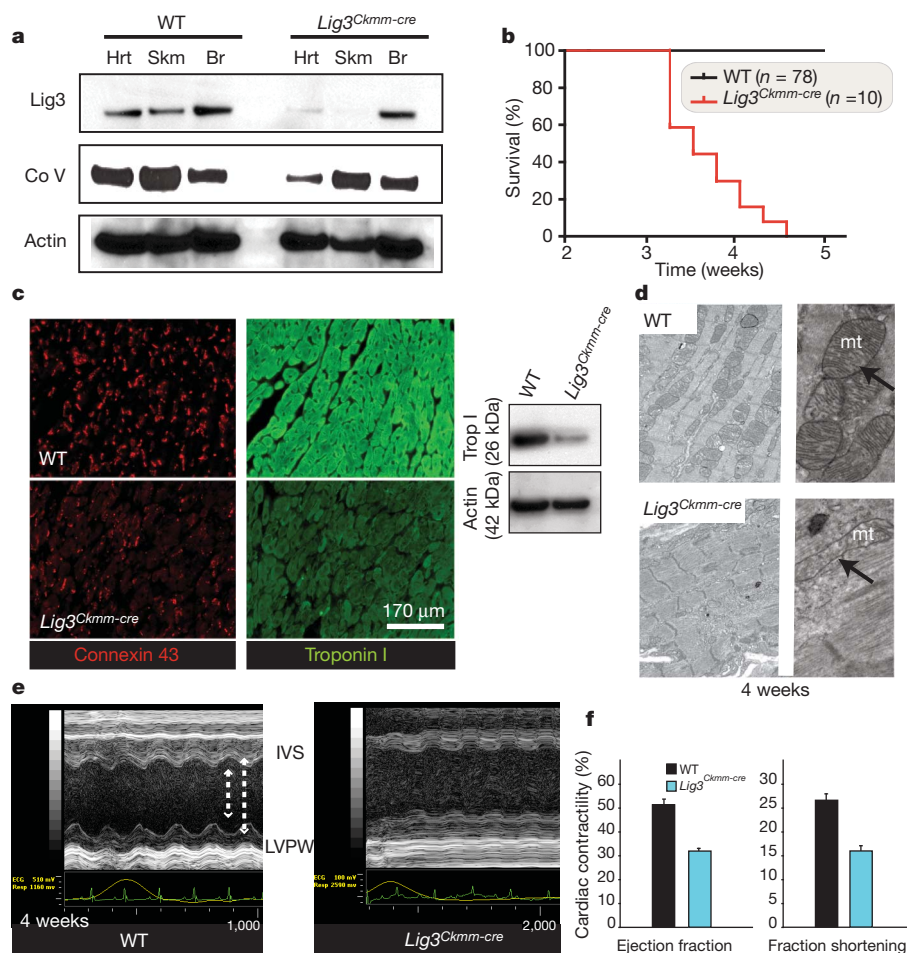


Figure 3 | *Lig3* inactivation causes cardiac failure associated with defective mitochondrial function. **a**, *Lig3* protein levels are reduced in *Lig3^{Ckmm-cre}* heart and skeletal muscle. A reduction of complex V subunit α also occurs in *Lig3^{Ckmm-cre}* heart. **b**, *Lig3^{Ckmm-cre}* mice die between 3.5 and 4.5 weeks of age; *n*, number of animals. **c**, Altered connexin-43 distribution and depletion of cardiac troponin I occurs in 4-week-old *Lig3^{Ckmm-cre}* heart. A western blot shows decreased cardiac troponin I levels. **d**, Electron microscopy of cardiac muscle shows disruption in myofibre structure and defects in mitochondrial

morphology. **e**, Echocardiographic analysis of *Lig3^{Ckmm-cre}* mice shows movement of the interventricular septum (IVS) and the posterior walls of the left ventricle (LVPW) were decreased in the *Lig3^{Ckmm-cre}* heart (arrows). Traces below the echocardiogram show ECG readings indicate abnormal heart rhythm. **f**, Cardiac contractility was also significantly depressed, with decreased ejection fraction and fraction shortening indicating that the pump function of the *Lig3^{Ckmm-cre}* heart is defective. Three individual animals of each genotype were used; error bars, standard deviation.

defects in short patch BER. This is the situation when another repair factor, aprataxin (Aptx), is defective²⁶. We used aphidicolin to inhibit acutely the long-patch-repair polymerases Pol- δ and Pol- ϵ , but found no DNA repair defect in *Lig3^{Nes-cre}* cells after addition of methyl methanesulphonate (Supplementary Fig. 8). In contrast to *Lig3^{Nes-cre}* cells, quiescent *Aptx^{-/-}* astrocytes had a reduced rate of DNA repair after addition of methyl methanesulphonate in the presence of aphidicolin, indicating long-patch-repair inhibition (Fig. 4d).

To further determine the repair capability and potential functional compensation by other DNA ligases in the absence of Lig3, we used lentiviral-mediated short hairpin RNA (shRNA) in primary mouse embryonic fibroblasts (MEFs) to inactivate Lig1, Lig3 and Lig4 (Fig. 4c). Importantly, under the conditions we established for shRNA-mediated ligase knockdown, minimal cell lethality was observed (data not shown). Using this approach, we found that, similar to *Lig3^{Nes-cre}* cells, Lig3-depleted cells were also competent to repair hydrogen-peroxide-induced DNA damage in quiescent cells (Fig. 4d). In contrast, inactivation of Lig1, but not Lig4, resulted in a marked defect in the ability to repair hydrogen-peroxide-induced DNA damage (Fig. 4d). Moreover, dual inactivation of Lig1 and Lig3 resulted in an exacerbated

repair deficiency that was similar to *Tdp1^{-/-}* cells, indicating functional cooperation between Lig1 and Lig3 after hydrogen peroxide treatment (Fig. 4d, e). We then compared other genotoxins including ionizing radiation and methyl methanesulphonate and found that, similar to damage induced by hydrogen peroxide, individual loss of Lig1, but not Lig3, strongly affected DNA repair (Fig. 4f and Supplementary Fig. 8). In the case of ionizing radiation, where both DNA single- and double-strand breaks occur, Lig4 inactivation also markedly reduced repair. Again, a cooperative defect was found after dual inactivation of Lig1 and Lig3 during repair of ionizing-radiation-induced lesions. The DNA repair defect associated with Lig1 inactivation was specific, as a modified Lig1 complementary DNA (cDNA) that was shRNA-resistant (*Lig1^{shR}*) restored repair capacity of the shLig1-targeted cells after treatment with hydrogen peroxide or methyl methanesulphonate (Fig. 4g and Supplementary Fig. 8).

Although *Xrcc1*-mediated stabilization of Lig3 in the nucleus suggests Lig3 is important during DNA repair^{2,7,8}, our data show that Lig3 is generally dispensable for *Xrcc1*-mediated DNA repair. On the other hand, our data show that Lig1 is clearly the operative ligase that repairs DNA lesions in quiescent cells usually ascribed to Lig3. However, it is

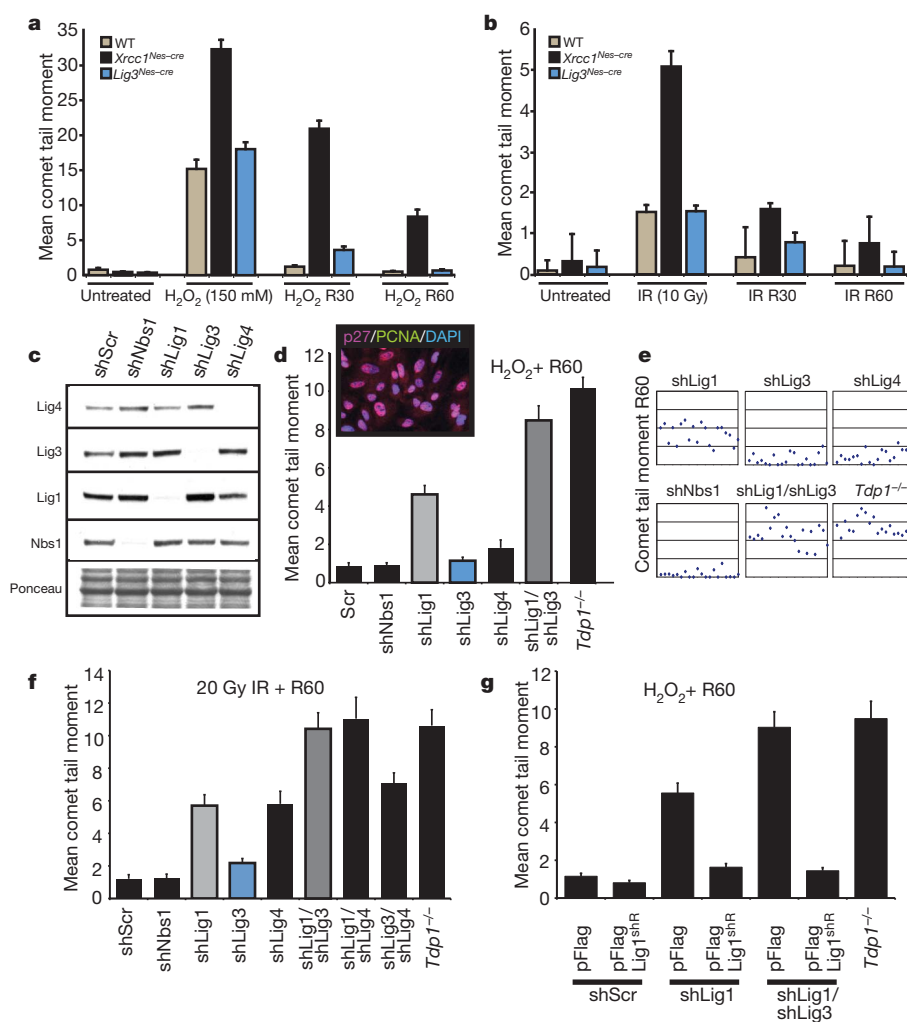


Figure 4 | Lig3 is not essential for nuclear DNA repair. **a, b**, Alkaline comet analysis shows DNA strand break repair after hydrogen peroxide treatment is similar between WT and *Lig3^{Nes-cre}* quiescent primary astrocytes, but is defective in *Xrcc1^{Nes-cre}* cells, as is repair after ionizing radiation (**b**); R30/R60 is recovery from the genotoxin after 30 or 60 min. Assays were run in triplicate from at least three separate experiments; error bars, s.e.m. **c**, Western blot analysis indicates specific shRNA-mediated knockdown of mammalian DNA ligases. Non-target scrambled (Scr) and Nbs1 shRNAs were included as controls, and Ponceau staining indicates relative protein transfer. **d**, Lig1 is

required for DNA repair after hydrogen peroxide damage; dual inactivation of Lig1 and Lig3 reveals cooperation between these ligases during repair. *Tdp1^{-/-}* MEFs are a repair-deficient control. Photograph indicates cells are quiescent based on p27⁺ and PCNA⁻ immunoreactivity. **e**, Representative individual comet tail moments of shRNA-mediated knockdowns after hydrogen peroxide treatment. **f**, Lig1 and Lig4 are required for DNA repair after ionizing radiation. Lig1 and Lig3 cooperate in the repair of ionizing-radiation-induced DNA damage. **g**, DNA repair competency after hydrogen peroxide treatment is restored using a shLig1-resistant version of Lig1 cDNA (*Lig1^{shR}*).

likely that the repair of specific subsets of DNA lesions do involve Lig3 given the increased repair deficiency after coordinate inactivation of Lig1 and Lig3. Although some controversy exists over the role of Lig1 during DNA repair^{27,28}, our data support Lig1 as a key player in this process. Increased sensitivity of repair assays may aid in further delineating Lig3 function during nuclear DNA repair.

The mitochondrial defects associated with Lig3 inactivation could occur at the level of mtDNA replication and/or repair. Generation of a mouse with a defect in the proofreading function of PolG led to accumulation of mtDNA mutations²⁰. However, although these animals were viable, they showed a variety of age-related defects associated with mtDNA mutations, and many survived longer than a year without developing the phenotypes we found after Lig3 inactivation in the brain or heart. In contrast, mitochondrial transcription factor A is essential for mtDNA replication, and its inactivation in the heart using *Ckmm-cre*²⁹ produced a similar phenotype to the *Lig3^{Ckmm-cre}* mice, suggesting defective mtDNA replication underpins the Lig3 mutant phenotype.

Although the Lig3-deficient nervous system (and heart) develops in a relatively normal manner, we assume mtDNA is nonetheless compromised during development of these tissues. The pronounced decline in the newborn *Lig3^{Nes-cre}* brain (and *Lig3^{Ckmm-cre}* heart) may be a combination of mtDNA defects acquired during organogenesis and stresses associated with the physiological and respiratory changes at birth³⁰, and the high demand for oxidative metabolism.

Thus, our data point to disparate biological roles for Lig3 and Xrcc1, with each primarily important in separate cellular compartments to ensure the integrity of mitochondrial and nuclear DNA, respectively. Additional analyses of Lig3 function during nuclear DNA repair and mtDNA maintenance will provide important insights into genome stability, mitochondrial biology and tissue homeostasis. Lig3 mutations may also underpin some of the many mitochondrial diseases for which the causative gene defect is unknown.

METHODS SUMMARY

Generation of Lig3 mutant mice. Embryonic stem cells were electroporated with a targeting vector incorporating *LoxP* sites flanking exons 6–14 of Lig3 to obtain mice with a floxed *Lig3* allele. Intercrossing these with *Nes-cre* or *Ckmm-cre* mice generated *Lig3^{Nes-cre}* or *Lig3^{Ckmm-cre}* animals.

Mitochondrial analysis. MitoTracker Red and PicoGreen were used to label astrocyte mitochondria. For electron microscopy, tissues were fixed in paraformaldehyde with osmium tetroxide post-fixing, and stained with uranyl acetate. mtDNA was assessed using pulsed-field gel electrophoresis and PCR.

Cardiac analysis. A VEVO-770 instrument (VisualSonics) was used for echocardiography analysis. ECG was determined and M-mode analysis was done to measure cardiac contractility.

DNA repair analysis. *Lig3^{Nes-cre}*, *Xrcc1^{Nes-cre}*, *Aptx^{-/-}* and control astrocytes were prepared as described previously⁶. MEFs were established from embryonic day (E) 13.5 embryos. RNA interference was used to inactivate DNA ligases in primary MEFs. shRNA-encoding lentiviral particles were produced by co-expression of pLKO.1-puro plasmids (TRC1, Sigma) with Mission Lentiviral packaging DNA (Sigma, SHP001) in 293T cells. The Comet Assay IV system (Perceptive Instruments) was used to assay DNA damage as described^{6,26}.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 28 April; accepted 22 December 2010.

- Ellenberger, T. & Tomkinson, A. E. Eukaryotic DNA ligases: structural and functional insights. *Annu. Rev. Biochem.* **77**, 313–338 (2008).
- Caldecott, K. W., McKeown, C. K., Tucker, J. D., Ljungquist, S. & Thompson, L. H. An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III. *Mol. Cell. Biol.* **14**, 68–76 (1994).
- Ljungquist, S., Kenne, K., Olsson, L. & Sandstrom, M. Altered DNA ligase III activity in the CHO EM9 mutant. *Mutat. Res.* **314**, 177–186 (1994).
- Lakshmipathy, U. & Campbell, C. Mitochondrial DNA ligase III function is independent of Xrcc1. *Nucleic Acids Res.* **28**, 3880–3886 (2000).
- Puebla-Osorio, N., Lacey, D. B., Alt, F. W. & Zhu, C. Early embryonic lethality due to targeted inactivation of DNA ligase III. *Mol. Cell. Biol.* **26**, 3935–3941 (2006).
- Lee, Y. *et al.* The genesis of cerebellar interneurons and the prevention of neural DNA damage require XRCC1. *Nature Neurosci.* **12**, 973–980 (2009).

- Almeida, K. H. & Sobol, R. W. A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification. *DNA Repair (Amst.)* **6**, 695–711 (2007).
- Caldecott, K. W. Single-strand break repair and genetic disease. *Nature Rev. Genet.* **9**, 619–631 (2008).
- Wang, H. *et al.* DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer Res.* **65**, 4020–4030 (2005).
- Lakshmipathy, U. & Campbell, C. The human DNA ligase III gene encodes nuclear and mitochondrial proteins. *Mol. Cell. Biol.* **19**, 3869–3876 (1999).
- Lakshmipathy, U. & Campbell, C. Antisense-mediated decrease in DNA ligase III expression results in reduced mitochondrial DNA integrity. *Nucleic Acids Res.* **29**, 668–676 (2001).
- Pinz, K. G. & Bogenhagen, D. F. Efficient repair of abasic sites in DNA by mitochondrial enzymes. *Mol. Cell. Biol.* **18**, 1257–1265 (1998).
- Krishnan, K. J. *et al.* What causes mitochondrial DNA deletions in human cells? *Nature Genet.* **40**, 275–279 (2008).
- LeDoux, S. P. *et al.* Repair of mitochondrial DNA after various types of DNA damage in Chinese hamster ovary cells. *Carcinogenesis* **13**, 1967–1973 (1992).
- Liu, P. *et al.* Removal of oxidative DNA damage via FEN1-dependent long-patch base excision repair in human cell mitochondria. *Mol. Cell. Biol.* **28**, 4975–4987 (2008).
- Tomkinson, A. E., Bonk, R. T. & Linn, S. Mitochondrial endonuclease activities specific for apurinic/apyrimidinic sites in DNA from mouse cells. *J. Biol. Chem.* **263**, 12532–12537 (1988).
- Van Houten, B., Woshner, V. & Santos, J. H. Role of mitochondrial DNA in toxic responses to oxidative stress. *DNA Repair (Amst.)* **5**, 145–152 (2006).
- Tynjismaa, H. & Suomalainen, A. Mouse models of mitochondrial DNA defects and their relevance for human disease. *EMBO Rep.* **10**, 137–143 (2009).
- Wallace, D. C. & Fan, W. The pathophysiology of mitochondrial disease as modeled in the mouse. *Genes Dev.* **23**, 1714–1736 (2009).
- Trifunovic, A. *et al.* Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* **429**, 417–423 (2004).
- Wallace, D. C., Fan, W. & Procaccio, V. Mitochondrial energetics and therapeutics. *Annu. Rev. Pathol.* **5**, 297–348 (2010).
- DiMauro, S. & Schon, E. A. Mitochondrial disorders in the nervous system. *Annu. Rev. Neurosci.* **31**, 91–123 (2008).
- Chan, D. C. Mitochondrial dynamics in disease. *N. Engl. J. Med.* **356**, 1707–1709 (2007).
- Chan, S. S. & Copeland, W. C. DNA polymerase gamma and mitochondrial disease: understanding the consequence of POLG mutations. *Biochim. Biophys. Acta* **1787**, 312–319 (2009).
- Moser, J. *et al.* Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase IIIα in a cell-cycle-specific manner. *Mol. Cell* **27**, 311–323 (2007).
- Reynolds, J. J. *et al.* Defective DNA ligation during short-patch single-strand break repair in ataxia oculomotor apraxia 1. *Mol. Cell. Biol.* **29**, 1354–1362 (2009).
- Bentley, D. J. *et al.* DNA ligase I null mouse cells show normal DNA repair activity but altered DNA replication and reduced genome stability. *J. Cell Sci.* **115**, 1551–1561 (2002).
- Sleeth, K. M., Robson, R. L. & Dianov, G. L. Exchangeability of mammalian DNA ligases between base excision repair pathways. *Biochemistry* **43**, 12924–12930 (2004).
- Wang, J. *et al.* Dilated cardiomyopathy and atrioventricular conduction blocks induced by heart-specific inactivation of mitochondrial DNA gene expression. *Nature Genet.* **21**, 133–137 (1999).
- Nakai, A. *et al.* Developmental changes in mitochondrial activity and energy metabolism in fetal and neonatal rat brain. *Brain Res. Dev. Brain Res.* **121**, 67–72 (2000).
- Simsek, D. *et al.* Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature* doi:10.1038/nature09794 (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Hartwell Center for biotechnology support, the Transgenic Core Facility for blastocyst injections and the Animal Resources Center for animal husbandry. We also acknowledge the Light Microscopy Core, the Electron Microscopy Core and C. Calabrese and the Animal Imaging Core for magnetic resonance imaging and echocardiography analysis. P.J.M. is supported by the National Institutes of Health (NS-37956, CA-21765), a Cancer Center Support Grant (P30 CA21765) and the American Lebanese and Syrian Associated Charities of St Jude Children's Research Hospital. S.K. is a Neoma Roadway AP Endowed Fellow.

Author Contributions Y.G., S.K. and Y.L. performed all experiments characterizing the Lig3-deficient mouse and contributed to writing the manuscript. Y.G. and H.R.R. generated the targeted embryonic stem cells for blastocyst injection and were responsible for colony production and maintenance with assistance from S.K. and Y.L. S.K. and J.Z. established the functional analyses of ligases. J.E.R. provided pathology analysis. P.J.M. was project leader and produced the final version of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.J.M. (peter.mckinnon@stjude.org).

METHODS

Ligase 3 mice. A targeting vector was designed to delete conditionally exons 6–14 using recombinase. A single *LoxP* site was inserted between exons 14 and 15 and a neomycin-thymidine kinase selection cassette flanked by *Frt* sites with a single *LoxP* site was inserted between exons 5 and 6. The targeting construct was transfected into embryonic stem cells by electroporation and selection with G418; homologous recombinants were screened by Southern blotting using both 3' and 5' probes outside the targeted region. Positive clones were transfected a second time with a *Frt* recombinase expressing vector. Cells were selected using flutridine to remove cells retaining the thymidine kinase and further screened by PCR using primers flanking the *LoxP* site or the *LoxP/Frt* remaining after recombination. Positive clones were blastocyst-injected to generate chimaeric mice. The male chimaeras were bred with C57BL/6 females to establish mice carrying the conditional *Lig3* allele. Further breeding with cre mice under the control of the Nestin promoter (JAX#003771: B6.Cg-Tg(Nes-cre)1Kln/J) resulted in deletion throughout the nervous system, whereas breeding with the *Ckmm2-cre* mice (JAX#006475: B6.FVB(129S4)-Tg(Ckmm2-cre)5Khn/J) deleted *Ligase 3* in skeletal and cardiac muscle.

Histology. Mice were perfused with 4% (w/v) buffered paraformaldehyde and cryoprotected in buffered 25% sucrose (w/v) solution. Brains were dissected, cut sagittally and sectioned at 10 μ m using an HM500M cryostat (Microm). Nissl staining used 1% (w/v) thionin. Haematoxylin and eosin staining was done according to standard procedures. Immunohistochemical and immunocytochemical staining of tissues and cells was performed with the antibodies listed below. For colorimetric visualization of positive signals, sections were incubated with antibodies overnight at room temperature after quenching endogenous peroxidase using 0.6% (v/v) hydrogen peroxide in methanol. Slides were washed with phosphate-buffered saline (PBS) three times, followed by incubation with biotinylated secondary antibody and avidin-biotin complex (Vectastain Elite kit, Vector Labs). Antibodies were used after citrate-buffer-based antigen retrieval. Immunoreactivity was visualized with the VIP substrate kit (Vector Labs) using the manufacturer's protocol. After staining, sections were counterstained with 0.1% (w/v) methyl green, dehydrated and mounted in DPX (Fluka). For fluorescent signals of immunoreactivity, FITC- or Cy3-conjugated secondary antibodies (Jackson Immunologicals) were used and counterstained with DAPI or propidium iodide (Vector Labs). For immunohistochemistry and immunocytochemistry, we used the following antibodies: calbindin (mouse, 1:2,000, Sigma), PCNA (mouse, 1:500, Santa Cruz Biotechnology), Pax2 (rabbit, 1:500, Zymed number 71-6000), CO III-core 2 (mouse, 1:2,000, MitoSciences), Ki67 (rabbit, 1:250, Vector Labs), COX IV (mouse, 1:500, Abcam), cardiac Troponin-I (mouse, 1:50 fluorescence and 1:500 western blots, Abcam) and connexin-43 (rabbit, 1:50, Cell Signaling). Apoptosis was detected using the TdT-mediated dUTP nick end labelling (TUNEL) assay with the ApopTag kit (Millipore) according to the manufacturer's protocol.

Western blot analysis. Western blot analysis was performed with tissues (cortex, cerebella, liver and spleen) both from P5 control mice (*Lig3^{LoxP/+};Nes-cre*) and from conditional knockout mice (*Lig3^{LoxP/LoxP};Nes-cre*). Protein extracts were prepared by using lysis buffer (50 mM Tris-HCl, 200 mM NaCl, 0.2% NP-40, 1% Tween-20 (v/v), 1 mM NaF, 1 mM sodium vanadate, 50 mM β -glycerophosphate, 2 mM PMSF, and Complete Protease Inhibitor (Roche) and quantified by Bradford assay (Bio-Rad). Proteins (50 μ g per lane) were separated through a 4–12% (w/v) Bis-Tris SDS polyacrylamide gel (Invitrogen) and transferred onto nitrocellulose membrane (Bio-Rad). Blots were sequentially immunostained with mouse anti-Lig3 antibody (1:500, BD Biosciences) followed by horseradish peroxidase-conjugated donkey anti-mouse (1:2,000, GE Healthcare) and detected using ECL Plus chemiluminescence reagent (GE Healthcare). To detect XRCC1 protein levels, the immunoblots were stained with rabbit anti-Xrcc1 antibody⁶ (1:1,000) and processed as described above. Immunoblots were also stained with mouse complex V subunit- α antibody (mouse, 1:1,000, MitoScience) and Pax2 (1:5,000, Abcam number 79389) and processed as described above. Actin antibody (1:1,000, Santa Cruz Biotech) was used as a protein-loading control.

Mitochondrial analysis. For mitochondria staining, primary astrocytes were incubated with prewarmed (37 °C) culture medium described below containing MitoTracker Red (1:200, Invitrogen) for 1 h under normal growth conditions followed by media replacement for an additional 20 min. For mtDNA staining, the astrocytes were incubated with prewarmed culture medium containing PicoGreen (3 μ l ml⁻¹, Invitrogen) for 1 h under normal growth conditions. Cells were imaged and processed using a Nikon C1Si confocal microscope and EZCI imaging software (Nikon USA) using a 561 DPSS laser (Melles Griot). Complex I activity assays were done with brain samples using a MitoSciences microplate assay kit (<http://www.mitosciences.com>) according to the manufacturer's directions. Cellular oxygen consumption was measured using an Oxygen Biosensor system 96-well plate (BD) according to the manufacturer's directions.

Mitochondria were isolated from brain tissue using the Qproteome™ Mitochondria Isolation Kit (QIAGEN) according to the manufacturer's directions, and mtDNA was isolated using the Mitochondrial DNA Isolation Kit (BioVision) according to the manufacturer's directions. PCR of mtDNA was done using the Expand Long Range, dNTPack (Roche) with the following conditions: initial denaturation at 92 °C for 30 s followed by annealing at 56 °C for 30 s and elongation at 68 °C for 60 s for 39 cycles, with a final 68 °C elongation for 10 min. Primers used were as follows: 'set A', forward, 5'-CCT TCA TCC TTC TCT CCC TAT GAG GA, reverse 5'-GGT TGT TTG ATC CTG TTT CGT GGA; 'set B', forward, 5'-CCC AGC TAC TAC CAT CAT TCA AGT, reverse 5'-CAG TAT GCT TAC CTT GTT ACG ACT.

Mitochondrial DNA isolated from equal amounts of *Lig3* mutant or control tissue was lysed in 1 mg ml⁻¹ proteinase K, 1% (w/v) *N*-lauroylsarcosine, (pH 9.5) in LM-agarose plugs overnight and subjected to pulsed-field gel electrophoresis. Mitochondrial DNA was separated for 24 h with gradual changes in pulses from 0.05 to 20 s at 3.0 V cm⁻¹ at an angle of 120° in 1% megabase agarose gel (Bio-Rad) with the CHEF Mapper Pulsed Field Electrophoresis system (Bio-Rad). DNA was visualized with ethidium bromide or Southern blot analysis using radioactive mtDNA as a probe.

Isolation of primary cells. Primary astrocytes were prepared from P2 mouse brains as described previously⁶. Cortices were dissociated by passage through a 5 ml pipette and cells were re-suspended in Dulbecco's modified Eagle's medium and Ham's nutrient mixture F-12 (1:1 DMEM/F12, Gibco-BRL) supplemented with 10% fetal bovine serum (v/v), 1 \times glutamax, 100 U ml⁻¹ penicillin, 100 μ g ml⁻¹ streptomycin and 20 ng ml⁻¹ epidermal growth factor (Millipore). Primary astrocytes were established in Primaria T-25 tissue culture flasks (Falcon) at 37 °C in a humidified CO₂-regulated (5%) incubator.

MEFs were prepared from E13.5 embryos. Embryos were minced using dissection scissors, passed through a 23-gauge syringe followed by a 70 μ m cell strainer. Cells were re-suspended in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (v/v), 1 \times glutamax, 100 U ml⁻¹ penicillin, 100 μ g ml⁻¹ streptomycin and β -mercaptoethanol and established in T-25 tissue culture flasks (Falcon) at 37 °C in a humidified CO₂-regulated (5%) incubator.

Magnetic resonance imaging. Two-week-old mutant and control mice (age and sex matched) were subjected to magnetic resonance imaging analysis. All animals were scanned on a 7 T Bruker Clinscan animal magnetic resonance imaging scanner (Bruker BioSpin MRI GmbH) as described previously⁶.

Electron microscopy. Mice were perfused with 4% paraformaldehyde, 100 mM cacodylate, 2.5% sucrose, (pH 7.4), and fixed overnight. Samples were post-fixed in osmium tetroxide, stained in 2% uranyl acetate, dehydrated and embedded in Epon Beem Capsules. Sections (80 nm *en face*) were prepared on a Leica UCT ultramicrotome and collected onto 200 mesh copper grids. Images were collected on Joel JEM-1200 EXII transmission electron microscope equipped with an AMTV600 digital camera.

DNA repair analysis. Quiescent primary astrocytes or MEFs (2×10^5 cells per millilitre per sample) were treated with either 150 μ M hydrogen peroxide for 10 min on ice, methylmethane sulphonate for 10 min at various concentrations at 37 °C or γ -irradiation (either 10 or 20 Gy, ¹³⁷Cs). After hydrogen peroxide and ionizing radiation treatments, cells were incubated for various times in drug-free medium at 37 °C. Cells were trypsinized and suspended in pre-chilled 1 \times PBS, mixed with an equal volume of 1.2% low-melting point agarose (Invitrogen) maintained at 42 °C and immediately layered onto frosted glass slides (Fisher) pre-coated with 0.6% agarose and maintained in the dark at 4 °C for all further steps. Slides were immersed in pre-chilled lysis buffer (2.5 M NaCl, 10 mM Tris-HCl, 100 mM EDTA (pH 8.0), 1% (v/v) Triton X-100, 3% (v/v) DMSO, pH 10) for 1 h, washed with pre-chilled distilled water (twice for 10 min each) and placed into pre-chilled alkaline electrophoresis buffer (50 mM NaOH, 1 mM EDTA, 1% DMSO) for 45 min. Electrophoresis was performed at 95 mA (astrocytes) or 190 mA (MEFs) for 25 min, followed by neutralization in 0.4 M Tris-HCl (pH 7.0). Comets were stained with SYBR Green (1:10,000 in 1 \times PBS, Sigma) for 10 min. A minimum of 100 comet tail moments were measured using the Comet Assay IV system (Perceptive Instruments) coupled to an Axioskop2 plus microscope (Zeiss) at $\times 200$ magnification.

Lentivirus particles were used to deliver and express shRNAs to knockdown murine Ligases 1, 3 and 4, with Nbs1 and a non-target scrambled shRNA used as controls. All MISSION shRNA plasmid DNA constructs were obtained from Sigma-Aldrich. The pLKO.1-puro parental vector expressed the TRC1 shRNAs and included the following: shLig1-5'-CCG GGC CCG GAC ATT TGA GAA GAT TCT CGA GAA TCT TCT CAA ATG TCC GGG CTT TTT G, shLig3-5'-CCG GCC ACC GAA CAG AAG CTC AAT ACT CGA GTA TTG AGC TTC TGT TCG GTG GTT TTT G and shLig4-5'-CCG GCC TCT CAG TAT GTT TCG ACA TCT CGA GAT GTC GAA ACA TAC TGA GAG GTT TTT G.

Lentiviral particles were made by co-transfection of pLKO.1-puro plasmids with Mission Lentiviral packaging DNA (Sigma, SHP001) into subconfluent 293T cells using Fugene 6 (Roche). Lentivirus was collected 48 h later and filtered (0.45 µm pore size) and primary MEFs (1.25×10^5 cells in 6 cm² plates grown for 20 h) were transduced with individual or a combination of lentiviral particles for 36 h at 37 °C with polybrene (10 µg ml⁻¹). All data reflect experiments performed in triplicate. Western blots were used to confirm shRNA knockdowns; antibodies used were Nbs1 (1/500, Cell Signaling), Lig1 (1/1,000, Santa Cruz Biotech), Lig3 (1/500, BD Biosciences) and Lig4 (1/400, Abcam).

Generation of an shLig1-resistant Lig1 cDNA (Lig1^{shR}) was done using the following primers: L1^{shR} forward-5'-CTT ACC TGG CTG TGG CAA GCA

CGT TCG AGA AAA TCG AGG AGG TGT CTG C and L1^{shR} reverse-5'-GCA GAC ACC TCC TCG ATT TTC TCG AAC GTG CTT GCC CAC GCC AGG TAA G. These silent mutations were introduced into human Lig1 cDNA using the Quikchange II Site-Directed Mutagenesis kit (Stratagene) according to the manufacturer's instructions. The modified Lig1 construct was co-transfected into MEFs with pmaxGFP plasmid (4:1 ratio) using Nucleofector methodology (Lonza) and then sorted by fluorescence-activated cell sorting to select for green fluorescent protein (GFP)-positive cells (~30%). Lig1^{shR}-transfected cells were then treated with shScr, shLig1 or a combination of shLig1/shLig3 lentivirus and assayed for DNA repair after genotoxic stress using the comet assay, as described above.

Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair

Deniz Simsek^{1,2}, Amy Furda³, Yankun Gao⁴, Jérôme Artus¹, Erika Brunet^{1,5,6,7}, Anna-Katerina Hadjantonakis^{1,2}, Bennett Van Houten³, Stewart Shuman^{2,8}, Peter J. McKinnon⁴ & Maria Jasin^{1,2}

Mammalian cells have three ATP-dependent DNA ligases, which are required for DNA replication and repair¹. Homologues of ligase I (Lig1) and ligase IV (Lig4) are ubiquitous in Eukarya, whereas ligase III (Lig3), which has nuclear and mitochondrial forms, appears to be restricted to vertebrates. Lig3 is implicated in various DNA repair pathways with its partner protein Xrcc1 (ref. 1). Deletion of *Lig3* results in early embryonic lethality in mice, as well as apparent cellular lethality², which has precluded definitive characterization of Lig3 function. Here we used pre-emptive complementation to determine the viability requirement for Lig3 in mammalian cells and its requirement in DNA repair. Various forms of Lig3 were introduced stably into mouse embryonic stem (mES) cells containing a conditional allele of *Lig3* that could be deleted with Cre recombinase. With this approach, we find that the mitochondrial, but not nuclear, Lig3 is required for cellular viability. Although the catalytic function of Lig3 is required, the zinc finger (ZnF) and BRCA1 carboxy (C)-terminal-related (BRCT) domains of Lig3 are not. Remarkably, the viability requirement for Lig3 can be circumvented by targeting Lig1 to the mitochondria or expressing *Chlorella* virus DNA ligase, the minimal eukaryal nick-sealing enzyme³, or *Escherichia coli* LigA, an NAD⁺-dependent ligase¹. *Lig3*-null cells are not sensitive to several DNA-damaging agents that sensitize Xrcc1-deficient cells^{4–6}. Our results establish a role for Lig3 in mitochondria, but distinguish it from its interacting protein Xrcc1.

Biochemical and cell biological experiments implicate the nuclear Lig3–Xrcc1 complex in single-strand break repair, short patch base excision repair and nucleotide excision repair¹. Lig3 and Xrcc1 interact through carboxy (C)-terminal BRCT domains found in each protein⁷. This interaction is important for the stability of Lig3 (ref. 7) and the recruitment of Lig3 to DNA damage foci⁸. Purified Lig3–Xrcc1 is proficient at nick sealing *in vitro*⁹, and the complex associates with several other proteins involved in single-strand break repair, including Parp1 (ref. 10), aprataxin and TDP1 (ref. 1).

Lig3 also has a mitochondrial form due to an alternative translation start site, which results in a mitochondrial leader sequence (MLS)¹¹. Mammals differ in this respect from budding yeast, where the Lig1 homologue, Cdc9, is the mitochondrial DNA ligase¹². In mitochondria, Lig3 appears to act independently of Xrcc1, as Xrcc1 is not present in this organelle¹³. Disruption of the *Lig3* gene, like *Xrcc1*, results in early embryonic lethality in the mouse^{2,5}, and *Lig3*-null cell lines could not be established from these animals². The similar timing of lethality of *Lig3* and *Xrcc1*-null embryos suggests that death could result from similar phenotypic consequences related to Lig3 nuclear functions in DNA repair. Alternatively, or in addition, the mitochondrial function of Lig3 may be critical for survival.

To determine whether *Lig3* is an essential gene because of its nuclear and/or mitochondrial function, we developed a pre-emptive

complementation strategy in mES cells (Fig. 1a). A *Lig3*^{KO/cKOneo+} cell line was constructed which contains one conditional *Lig3* allele with an intronic *Neo* selection marker and *LoxP* sites flanking exons 6 and 14 and a second allele in which these exons were already removed by Cre recombinase (Fig. 1a and Supplementary Fig. 1). These exons encode part of the DNA binding domain and the catalytic core of the protein. Cre recombinase was expressed in the *Lig3*^{KO/cKOneo+} cells, and 145 clones were replica-plated in media with or without G418. No G418-sensitive clones (that is, *Lig3*^{KO/KO}) were obtained (Fig. 1b), consistent with the requirement for Lig3 for cellular viability. We then stably integrated transgenes expressing wild-type (WT), mitochondrial or nuclear Lig3; the nuclear (NucLig3) version lacked the MLS, and the mitochondrial (MtLig3) version contained the MLS but was mutated at the nuclear translation initiation site (M88T) (Fig. 1c and Supplementary Fig. 2). Green fluorescent protein (GFP) fusions of these proteins were also tested (Supplementary Fig. 3a).

To determine which *Lig3* transgenes allow the survival of cells deleted for endogenous *Lig3*, Cre recombinase was used to transform *Lig3*^{KO/cKOneo+} cells to *Lig3*^{KO/KO} cells. A large fraction of the post-Cre clones expressing WT Lig3 or MtLig3 were G418 sensitive (34–50%), whereas no G418-sensitive clones were obtained with NucLig3 (Fig. 1b and Supplementary Fig. 4). We confirmed that G418-sensitive cells were *Lig3*^{KO/KO} (Fig. 1a) and that endogenous Lig3 was no longer present, with the only Lig3 present in the cells expressed from the transgene (Fig. 1d). Thus, cellular viability requires mitochondrial Lig3. To determine whether DNA ligase activity was essential for cell survival, we introduced a K508V mutation that abolishes ligase adenylation and nick-sealing¹ into MtLig3. No G418-sensitive clones were derived from four independent transgenic cell lines (Fig. 1b), demonstrating that the requirement for mitochondrial Lig3 depends on its ligase activity.

BRCT domains are frequently involved in protein–protein interactions, and the BRCT domain of Lig3 is known to interact with Xrcc1 (ref. 7). However, as Xrcc1 is not found in mitochondria¹³, the role of the BRCT domain for mitochondrial function of Lig3 is uncertain. Loss of the BRCT domain had no effect on the presence of Lig3 in mitochondria (Lig3-ΔBRCT and MtLig3-ΔBRCT; Supplementary Fig. 3a and data not shown). *Lig3*^{KO/KO} clones expressing Lig3-ΔBRCT or MtLig3-ΔBRCT (Fig. 2a) were recovered as a substantial fraction of clones (39–49%; Fig. 2b), indicating that the BRCT domain is not required for viability. Thus, MtLig3 does not have a partner protein bound to its BRCT domain that is essential for its function.

A unique feature of Lig3 compared with the other mammalian DNA ligases is a ZnF at its amino (N) terminus. The Lig3 ZnF interacts with Parp1 (ref. 10), and this interaction is reported to be important for the association of Lig3 with mitochondrial DNA (mtDNA)¹⁴. Biochemical studies have shown that the ZnF promotes DNA nick recognition¹⁵ and

¹Developmental Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. ²Weill Cornell Graduate School of Medical Sciences, New York, New York 10065, USA.

³Department of Pharmacology and Chemical Biology, University of Pittsburgh School of Medicine and The University of Pittsburgh Cancer Institute, Hillman Cancer Center, Pittsburgh, Pennsylvania 15213, USA. ⁴Department of Genetics and Tumor Cell Biology, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. ⁵Museum National d'Histoire Naturelle, 43 rue Cuvier, F-75005 Paris, France. ⁶CNRS, UMR7196, 43 rue Cuvier, F-75005 Paris, France. ⁷INSERM, U565, 43 rue Cuvier, F-75005 Paris, France. ⁸Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA.

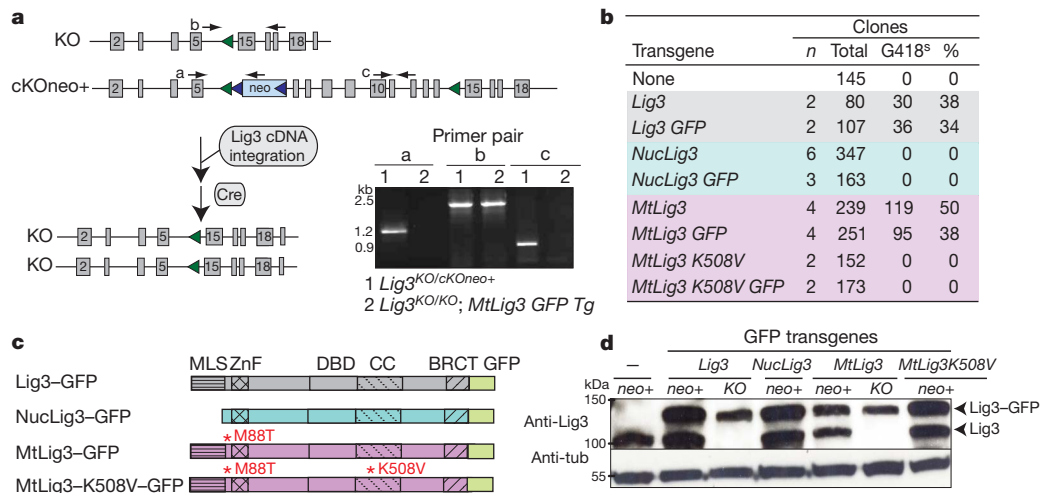


Figure 1 | Mitochondrial Lig3 activity is critical for cellular viability. **a**, Pre-emptive complementation strategy for *Lig3* deletion. *Lig3*^{KO/KO} clones were identified by their lack of growth in G418 (neo⁻). PCR confirmed the genotype, as indicated. KO, knockout; Tg, transgene. **b**, Mitochondrial Lig3 activity is critical for cellular viability. *n*, number of independently derived *Lig3* transgenic cell lines analysed; Total, total number of post-Cre colonies analysed by replica plating with and without G418; G418^S, G418 sensitive; %, ratio of

intermolecular ligation¹⁶. Nonetheless, *Lig3*^{KO/KO} clones expressing *Lig3*-ΔZnF or *MtLig3*-ΔZnF (Fig. 2a) were efficiently recovered after Cre expression (37–38%; Fig. 2b).

Our results reveal that the catalytic activity of Lig3 is critical for cell survival, but that the ZnF and BRCT domains, which interact with various proteins, are dispensable, raising the question whether Lig3 itself is critical for mitochondrial function, or whether another DNA ligase would substitute. As the Lig1 homologue in yeast provides

G418^S to Total. **c**, *Lig3* proteins tested for pre-emptive complementation. DBD, DNA-binding domain; CC, catalytic core; *M88T, mutation of the nuclear translation initiation site. **d**, Western blot analysis showing the loss of endogenous Lig3 in *Lig3*^{KO/KO} clones stably expressing GFP-tagged Lig3 or *MtLig3*. Lig3 is 105 kDa, whereas the GFP fusions are approximately 135 kDa. neo+, *Lig3*^{KO/KO}; KO, *Lig3*^{KO/KO}; α-tub, α-tubulin.

mitochondrial ligase function¹², we provided an MLS to murine Lig1 (Fig. 2a). GFP-tagged *MtLig1*, but not WT Lig1, was localized to mitochondria (Supplementary Fig. 3c), as expected. Stable *Lig3*^{KO/KO} cell lines expressing *MtLig1*, but not WT Lig1, could be efficiently converted to *Lig3*^{KO/KO} (35%; Fig. 2b). *MtLig1* *Lig3*^{KO/KO} clones were devoid of Lig3, and expressed instead a larger Lig1 protein due to the GFP tag (Fig. 3a). Thus, targeting Lig1 to mitochondria circumvents the viability requirement for Lig3, allowing the creation of *Lig3*-null cells. In this way, the DNA ligase repertoire of mammalian cells is converted to that of yeast.

Given that ZnF and BRCT-truncated forms of Lig3 and *MtLig1* could rescue *Lig3*^{KO/KO} cells, we investigated their proficiency in mtDNA maintenance and repair. The mtDNA copy number was maintained as well (or better) in these cell lines than in WT *Lig3*-expressing cells (Fig. 2c), indicating that cells expressing these altered ligases are competent to replicate their mtDNA during continued passage. A long-range quantitative PCR (qPCR) assay¹⁷ was performed

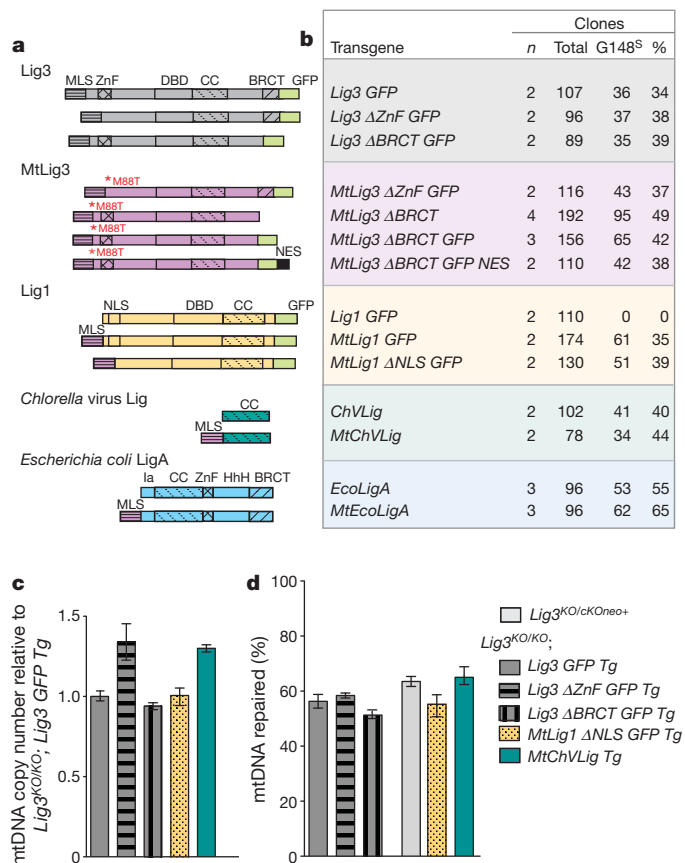


Figure 2 | Mitochondrial DNA ligase activity can be provided by a variety of DNA ligases. **a**, DNA ligases tested for pre-emptive complementation of *Lig3*^{KO/KO} cells. ΔZnF, deletion of Lig3 amino acids 89–258; ΔBRCT, deletion of Lig3 amino acids 934–1009; ΔNLS, deletion of Lig1 amino acids 135–147; NES, nuclear export signal from MAPKK²¹; Mt, presence of Lig3 MLS; NLS, nuclear localization signal; HhH, helix–hairpin–helix. **b**, Mitochondrial DNA ligase activity can be provided by a variety of DNA ligases. ChV Lig and EcoLigA presumably enter mitochondria without the requirement for an MLS. **c**, Cells expressing exogenous DNA ligases are competent to replicate and maintain mtDNA. mtDNA copy number was quantified by qPCR by amplifying a 117 base pair (bp) fragment from mtDNA. Values are presented relative to levels of mtDNA in *Lig3*^{KO/KO}; *Lig3* GFP Tg cells. Data represent the mean of two biological repeats each determined twice by qPCR ± s.e.m. **d**, Cells expressing exogenous DNA ligases showed similar capacities for mtDNA repair after oxidative damage. Cells were treated with 175 μM hydrogen peroxide for 15 min and allowed to recover for 1.5 h. To measure repair, a 10-kb mtDNA fragment was amplified after damage and quantified by qPCR. Values were normalized to the amplification of a 117-bp mtDNA fragment. Percentage repair is 1 minus the amount of damage remaining after 1.5 h recovery divided by the initial damage. There was no significant difference between cell lines expressing WT Lig3 (parental cells and transgene-rescued cells) and the other ligase forms. For cells expressing *MtLig1*-ΔNLS and *MtChV Lig*, two transgenic cell lines were analysed. Data represent the mean of two to four determinations on multiple clones with each qPCR performed twice ± s.e.m.

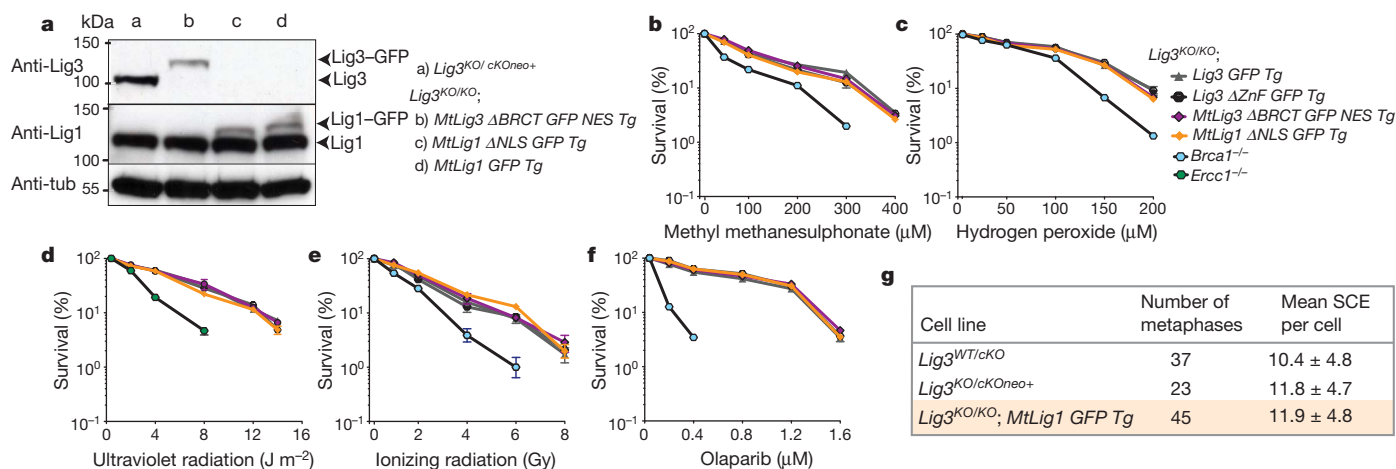


Figure 3 | Loss of Lig3 is not associated with sensitivity to several DNA-damaging agents or with increased SCE. **a**, Western blot analysis showing the loss of endogenous Lig3 protein in *Lig3^{KO/KO}* cells with the indicated transgenes. **b–f**, Sensitivity of *Lig3^{KO/KO}* cell lines to the indicated DNA-damaging agents was measured using colony formation assays. *Brca1^{-/-}* and *Erc1^{-/-}* cells are only shown on graphs when they are sensitive. For each cell

line and agent, $n = 4$ and error bars show s.e.m.; error bars in some cases are smaller than the symbol. **g**, SCE analysis. The range of SCEs was between 5 and 21 per metaphase for each cell line. For cells expressing MtLig1, two transgenic cell lines were analysed. The differences between the cell lines are not significant using a two-tailed unpaired t -test. Values are presented with one s.d. from the mean.

to assess the mitochondrial base excision repair capacity of these cells in response to oxidative damage, and these altered ligases were similarly proficient in repairing mtDNA lesions compared with WT Lig3-expressing cells (Fig. 2d).

At 298 amino acids, *Chlorella* virus DNA ligase (ChVLIg) is the smallest eukaryal ligase known, consisting solely of a catalytic core³. We expressed ChVLIg and a modified form containing an MLS, MtChVLIg (Fig. 2a), and found that expression of either allowed deletion of *Lig3* from the mouse genome (Fig. 2b). It is conceivable that ChVLIg contains an internal sequence that allows translocation into mitochondria¹⁸. Thus, a minimal ATP-dependent ligase, devoid of auxiliary domains, rescues the survival of *Lig3*-null mammalian cells. Further, *Lig3*-null cells rescued by MtChVLIg were proficient at mtDNA maintenance (Fig. 2c) and repair (Fig. 2d).

Whereas ATP-dependent ligases are widespread, ligases that use NAD⁺ as a cofactor are usually restricted to bacteria¹. *E. coli* DNA ligase, LigA, is NAD⁺ dependent and has a distinctive domain organization compared with mammalian ligases¹. LigA and a modified form with an MLS (Fig. 2a) were expressed from transgenes, and like ChVLIg, both forms were found to allow the survival of *Lig3*-null cells (Fig. 2b). Hence, there is no essential functional distinction between NAD⁺ and ATP-dependent ligases in the mammalian mitochondria, akin to swaps of NAD⁺ and ATP-dependent ligases performed in bacteria¹⁹ and yeast²⁰.

We demonstrated that nuclear Lig3 is not required for cell survival, as MtLig1 *Lig3^{KO/KO}* cells are null for *Lig3*. To impair nuclear localization of MtLig1, we removed the Lig1 nuclear localization signal, creating MtLig1-ΔNLS *Lig3^{KO/KO}* cells (Fig. 2a, b and Supplementary Fig. 3c). MtLig1-ΔNLS, like MtLig1, was expressed at a substantially lower level than endogenous Lig1 (Fig. 3a). As a complementary approach, we also created *Lig3^{KO/KO}* cells expressing MtLig3-ΔBRCT-NES, whose interaction with Xrcc1 is abrogated and which is excluded from the nucleus by addition of a potent nuclear export signal (NES)²¹ (Figs 2a, b and 3a and Supplementary Fig. 3b).

To assess the nuclear role of Lig3, we tested the sensitivity of *Lig3*-null (*Lig3^{KO/KO}*; MtLig1-ΔNLS) and nuclear Lig3-deficient (*Lig3^{KO/KO}*; MtLig3-ΔBRCT-NES) cells to a variety of DNA-damaging agents. Xrcc1-deficient cells are highly sensitive to alkylating agents like methyl methanesulphonate^{4–6}. If the interaction of Xrcc1 with Lig3 is relevant to base excision repair, cells without nuclear Lig3 would also be expected to be sensitive to methyl methanesulphonate; however, we found that these cells were no more sensitive than transgenic cells

expressing WT Lig3 (Fig. 3b) or the parental cells (Supplementary Fig. 5). Xrcc1-deficient cells are also sensitive to agents that produce DNA single- and double-strand breaks, including hydrogen peroxide and ionizing radiation^{4–6}, and to ultraviolet radiation²². By contrast, we found that cells without nuclear Lig3 were not any more sensitive to these agents than control cells (Fig. 3c–e and Supplementary Fig. 5). Thus, Lig3 appears to be dispensable for nuclear DNA damage repair that requires Xrcc1. Finally, we tested sensitivity to Parp inhibition, which causes the accumulation of single-strand breaks²³, and found that nuclear Lig3 was also not required for resistance of cells to Parp inhibition (Fig. 3f).

As the ZnF domain of Lig3 has been reported to be critical for its intermolecular ligation activity¹⁶, we also investigated whether deletion of this domain in the context of an otherwise WT Lig3 would impair resistance of cells to ionizing radiation. As with the other mutants, *Lig3*-ΔZnF *Lig3^{KO/KO}* cells were no more sensitive than control cells (Fig. 3e).

Xrcc1-deficient cells are notable for their high rate of spontaneous sister-chromatid exchange (SCE): both mouse and hamster Xrcc1 mutants have approximately tenfold higher SCE levels than control cells^{4,5}. We examined spontaneous SCEs in MtLig1 *Lig3^{KO/KO}* cells and found levels similar to control cells (Fig. 3g). Thus, the high level of SCEs found with Xrcc1 deficiency is not recapitulated with Lig3 deficiency.

The lack of Lig3 in many model organisms has limited their use to study its function. In mouse, disruption of any of the DNA ligase genes leads to embryonic death, but the most severe phenotype occurs with Lig3 disruption^{2,24,25}. Lig1 has been considered to be the replicative ligase^{1,26}, but the earlier death associated with Lig3 disruption, together with the inability to obtain viable Lig3-null cells, left open the possibility that Lig3 could have a critical role in nuclear DNA metabolism. The generation of viable and healthy *Lig3*-null cells by providing a mitochondrial ligase conclusively rules out an essential role for Lig3 in the nucleus.

The well-documented interaction between Lig3 and Xrcc1 had suggested that Lig3 would be critical for the same nuclear DNA repair pathways as Xrcc1, similar to the Lig4–Xrcc4 complex in DSB repair¹. However, the lack of sensitivity of Lig3-null cells to the spectrum of DNA-damaging agents that sensitize Xrcc1-deficient cells, as shown here in proliferating cells and in the accompanying paper in quiescent cells²⁷, together with a normal SCE frequency, provides strong evidence that Lig3 is not required for Xrcc1-dependent nuclear DNA repair, pointing instead to a role for Lig1.

Our results demonstrate instead that *Lig3* is an essential gene because of its requirement in mitochondria. However, Lig3 can be replaced in mitochondria with Lig1, the mitochondrial ligase in lower eukaryotes, with an algal viral ligase consisting solely of a catalytic core, and even the NAD⁺-dependent *E. coli* LigA. Thus, these results attest to the requirement for a functional DNA ligase, which trumps even cofactor specificity. Why vertebrates developed a requirement for Lig3 is uncertain, but given our results, the additional domains found in Lig3 do not appear to be essential for mitochondrial function, including mtDNA maintenance or repair of oxidative damage. These results emphasize a surprising plasticity that mammalian cells have in their mitochondrial DNA ligase requirement.

METHODS SUMMARY

Cell culture. To construct stable cell lines expressing various DNA ligases, 5×10^6 *Lig3^{KO/cKOneo+}* cells were electroporated with 12 µg ligase expression vector at 800 V, 3 µF. Hygromycin-resistant clones were picked after incubation for 10 days in 150 µg µl⁻¹ hygromycin. Initial screening for exogenous ligase expression was performed by PCR with reverse transcription (RT-PCR) using specific primers, followed by western blotting. For deletion of the endogenous *Lig3* allele, 5×10^6 cells were electroporated with 5 µg Cre recombinase vector at 250 V, 950 µF. Cells were plated based on a serial dilution. After 7 days, colonies were picked and expanded, and then replica plated into two 96-well plates. One plate was cultured with 200 µg µl⁻¹ G418, whereas the other plate was cultured in normal media. Clones that did not grow in G418, but grew in normal media, have converted the *Lig3^{KO/cKOneo+}* allele to a *Lig3^{KO}* allele. The genotype of these clones was confirmed by PCR.

Western blotting. Whole-cell extracts were prepared with Nonidet-P40 buffer and were run on a 7.5% (w/v) Tris-HCl SDS page gel, blotted and then probed with Lig3 antibody clone 7 (BD Transduction Labs), which recognizes both the human and mouse Lig3 proteins, or Lig1 antibody N-13 (Santa Cruz). α -Tubulin (Sigma) was used as a loading control.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 April 2010; accepted 5 January 2011.

- Ellenberger, T. & Tomkinson, A. E. Eukaryotic DNA ligases: structural and functional insights. *Annu. Rev. Biochem.* **77**, 313–338 (2008).
- Puebla-Osorio, N., Lacey, D. B., Alt, F. W. & Zhu, C. Early embryonic lethality due to targeted inactivation of DNA ligase III. *Mol. Cell. Biol.* **26**, 3935–3941 (2006).
- Ho, C. K., Van Etten, J. L. & Shuman, S. Characterization of an ATP-dependent DNA ligase encoded by *Chlorella* virus PBCV-1. *J. Virol.* **71**, 1931–1937 (1997).
- Thompson, L. H., Brookman, K. W., Jones, N. J., Allen, S. A. & Carrano, A. V. Molecular cloning of the human XRCC1 gene, which corrects defective DNA strand break repair and sister chromatid exchange. *Mol. Cell. Biol.* **10**, 6160–6171 (1990).
- Tebbs, R. S. *et al.* Requirement for the Xrcc1 DNA base excision repair gene during early mouse development. *Dev. Biol.* **208**, 513–529 (1999).
- Lee, Y. *et al.* The genesis of cerebellar interneurons and the prevention of neural DNA damage require XRCC1. *Nature Neurosci.* **12**, 973–980 (2009).
- Caldecott, K. W., McKeown, C. K., Tucker, J. D., Ljungquist, S. & Thompson, L. H. An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III. *Mol. Cell. Biol.* **14**, 68–76 (1994).
- Mortusewicz, O., Rothbauer, U., Cardoso, M. C. & Leonhardt, H. Differential recruitment of DNA ligase I and III to DNA repair sites. *Nucleic Acids Res.* **34**, 3523–3532 (2006).
- Chen, X. *et al.* Distinct kinetics of human DNA ligases I, III α , III β , and IV reveal direct DNA sensing ability and differential physiological functions in DNA repair. *DNA Repair (Amst.)* **8**, 961–968 (2009).
- Leppard, J. B., Dong, Z., Mackey, Z. B. & Tomkinson, A. E. Physical and functional interaction between DNA ligase III α and poly(ADP-ribose) polymerase 1 in DNA single-strand break repair. *Mol. Cell. Biol.* **23**, 5919–5927 (2003).

- Lakshminpathy, U. & Campbell, C. The human DNA ligase III gene encodes nuclear and mitochondrial proteins. *Mol. Cell. Biol.* **19**, 3869–3876 (1999).
- Willer, M., Rainey, M., Pullen, T. & Stirling, C. J. The yeast CDC9 gene encodes both a nuclear and a mitochondrial form of DNA ligase I. *Curr. Biol.* **9**, 1085–1094 (1999).
- Lakshminpathy, U. & Campbell, C. Mitochondrial DNA ligase III function is independent of Xrcc1. *Nucleic Acids Res.* **28**, 3880–3886 (2000).
- Rossi, M. N. *et al.* Mitochondrial localization of PARP-1 requires interaction with mitofilin and is involved in the maintenance of mitochondrial DNA integrity. *J. Biol. Chem.* **284**, 31616–31624 (2009).
- Mackey, Z. B. *et al.* DNA ligase III is recruited to DNA strand breaks by a zinc finger motif homologous to that of poly(ADP-ribose) polymerase. Identification of two functionally distinct DNA binding regions within DNA ligase III. *J. Biol. Chem.* **274**, 21679–21687 (1999).
- Cotner-Gohara, E., Kim, I. K., Tomkinson, A. E. & Ellenberger, T. Two DNA-binding and nick recognition modules in human DNA ligase III. *J. Biol. Chem.* **283**, 10764–10772 (2008).
- Santos, J. H., Meyer, J. N., Mandavilli, B. S. & Van Houten, B. Quantitative PCR-based measurement of nuclear and mitochondrial DNA damage and repair in mammalian cells. *Methods Mol. Biol.* **314**, 183–199 (2006).
- Neupert, W. & Herrmann, J. M. Translocation of proteins into mitochondria. *Annu. Rev. Biochem.* **76**, 723–749 (2007).
- Park, U. E., Olivera, B. M., Hughes, K. T., Roth, J. R. & Hillyard, D. R. DNA ligase and the pyridine nucleotide cycle in *Salmonella typhimurium*. *J. Bacteriol.* **171**, 2173–2180 (1989).
- Sriskanda, V., Schwer, B., Ho, C. K. & Shuman, S. Mutational analysis of *Escherichia coli* DNA ligase identifies amino acids required for nick-ligation *in vitro* and for *in vivo* complementation of the growth of yeast cells deleted for CDC9 and LIG4. *Nucleic Acids Res.* **27**, 3953–3963 (1999).
- Henderson, B. R. & Eleftheriou, A. A comparison of the activity, sequence specificity, and CRM1-dependence of different nuclear export signals. *Exp. Cell Res.* **256**, 213–224 (2000).
- Moser, J. *et al.* Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III α in a cell-cycle-specific manner. *Mol. Cell* **27**, 311–323 (2007).
- Farmer, H. *et al.* Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005).
- Bentley, D. *et al.* DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability. *Nature Genet.* **13**, 489–491 (1996).
- Frank, K. M. *et al.* Late embryonic lethality and impaired (V)DJ recombination in mice lacking DNA ligase IV. *Nature* **396**, 173–177 (1998).
- Barnes, D. E., Tomkinson, A. E., Lehmann, A. R., Webster, A. D. & Lindahl, T. Mutations in the DNA ligase I gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. *Cell* **69**, 495–503 (1992).
- Gao, Y. *et al.* DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair. *Nature* doi:10.1038/nature09773 (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Caldecott for the gift of the Lig3 expression vector, and M. Sanz for initial assistance with the SCE analysis. We also thank the members of Jasin laboratory, especially Y. Akamatsu, J. LaRocque, E. Kass and F. Vanoli, for discussions. This work was supported by PA CURE (to B.V.H.) and by National Institutes of Health grants ES019566 (to B.V.H.), NS37956 and CA21765 (to P.J.M.), and GM54668 (to M.J.).

Author Contributions D.S. performed most of the experiments. D.S. and M.J. designed the research and wrote the paper. A.F. performed the long-range qPCR assays to investigate mitochondrial BER and mitochondrial DNA maintenance, and with B.V.H. analysed the data. Y.G. and P.J.M. designed the initial Lig3 targeting scheme and generated the *Lig3^{wt/cKO}* embryonic stem cells. J.A. and A.-K.H. acquired confocal images for GFP-tagged proteins. E.B. contributed technical assistance and preparation of the manuscript. S.S. contributed discussions, provided reagents and shared unpublished data.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.J. (m-jasin@ski.mskcc.org).

METHODS

DNA constructs. A vector containing WT human Lig3 complementary DNA (cDNA) (with both mitochondrial and nuclear translation initiation sites), a gift from K.W. Caldecott, was digested with NheI and XbaI and subcloned into the NheI site of pCAGGS. As the cDNA contained a 51-bp linker located before the nuclear translation initiation site, it was modified by site-directed mutagenesis to remove the linker, with the primers 5'-GTGGCCCTGTGAGATGGCTGAGCA ACG-3' and 5'-CGTTGCTCAGCCATCTCACAGGGGCCAC-3' to restore an unmodified Lig3 sequence, creating pCAGGS-Lig3. A Pkg-hygromycin resistance gene was inserted at the PsiI site to create pCAGGS-Lig3-hyg. MtLig3 was generated by using site-directed mutagenesis to generate an M88T mutation in pCAGGS-Lig3-hyg using the primers 5'-GAGAGGCCCTGTGAGACCGCTGAGCA-3' and 5'-GAGAGGATCCCTAGCAGGGAGCTACCACTCTC-3'. For NucLig3, amino acids 1–87 were deleted by introducing NotI and BamHI sites into pCAGGS-Lig3-hyg by PCR using the primers 5'-GCATGCGGCCGCTGTGAG ATGGCTGAGCAACGGT-3', 5'-GGATGGATCCCTAGCAGGGAGCTACCA GTC-3'. For the ΔBRCT mutation, amino acids 934–1009 were deleted by introducing NheI and MfeI sites by PCR using the primers 5'-GGCCGCTAGCG GGCAGCTATATGCTTTGGCTTCAAGAT-3' and 5'-GAGACAATTGTTA CTATACCTTTGTTTGGCACAGCGTC-3'. The ΔZnF mutation was generated by deleting amino acids 89–258 using site-directed mutagenesis with primers 5'-TGGCCCTGTGAGATGAAGGACTGTCTGCTAC-3' and 5'-GTAGCAG ACAGTCTTCTATCTACAGGGGCCA-3'. For GFP tagging of the Mt-tagged constructs, SacII and AgeI sites were introduced and stop codons of the full-length or ΔBRCT proteins were converted into alanine codons by PCR and cloned in frame into SacII and AgeI sites of pEGFP-N1 (Clontech). PCR primers for full length were 5'-ACGGTACCGCGGCAGCTATATGCTTTGG-3' and 5'-ACGG TACCGCGGCAGCTATATGCTTTGG-3', and for ΔBRCT were 5'-ACGGTAC CGCGGCAGCTATATGCTTTGG-3' and 5'-GGCGACCGGTGGTACCTTT GTTTGGCACAGCG-3'.

For other constructs with GFP fusions (NucLig3, Lig3, ΔZnF and K508V), plasmids were digested with PmlI and ligated into the vector backbone of MtLig3-GFP using the same enzyme. The MAPKK NES²¹ was fused to the C terminus of GFP by PCR using the primers 5'-GCCCCCTACGCCAGTACC AAGAA-3' and 5'-GGCCAATTGGCCTTATTACTGCTGCTCGTCCAGCTC CAGCTCCTCCAGCTTCTTTGGAGGTCCACGAGATTCTTGTACAGCTCG TCCAT-3'. Mouse Lig1 cDNA (Invitrogen) was amplified with primers introducing KpnI and AgeI sites and changing a stop codon into an alanine codon; this fragment was cloned in frame into the KpnI and AgeI sites of pEGFP-N1. The Lig3 MLS was amplified with the primers 5'-GGCGAATTCTATATGCTTTGGC TTCAAGATCTTCTTTC-3' and 5'-ATTGGTACCCCTCACAGGGGCCACTG CAG-3' and cloned into the EcoRI and KpnI sites of Lig1-GFP-hyg vector. The *Chlorella* virus DNA ligase coding region was amplified with ChV-NheI and ChV-MfeI primers and cloned into the NheI and MfeI sites of the pCAGGS-Lig1-Hyg vector. ChV-NheI: 5'-GCCGCTAGCACCATGGCAATCACAAAGCCATT-3'; ChV-MfeI: 5'-GCCCAATTGTTAACGGTCTTCTCGTGAC-3'. The *E. coli* DNA ligase coding region was amplified with LigA-NheI and LigA-MfeI primers and cloned into the NheI and MfeI sites of the pCAGGS-Lig1-Hyg vector. LigA-NheI: 5'-GCCGCTAGCACCATGGCAATCAATCGAACAACAA-3'; LigA-MfeI: 5'-GCCCAATTGTCAGCTACCCAGCAACG-3'.

RT-PCR. Hygromycin-resistant clones were screened by RT-PCR using primers specific to human Lig3. A primer pair was used with the forward primer to the pCAGGS backbone and the reverse primer to exon 3 of human Lig3, resulting in a size difference for mitochondrial and nuclear forms (Supplementary Fig. 3): pCAGRTfw 5'-CAACGTGCTGGTTATTGTGC-3', hLig3Rv 5'-ACAGCTTTC TTCTTTGGTGTAACCT-3'. A similar strategy was used for Lig1, with primers pCAGRTfw and Lig1RT_RV (5'-ACCGCTGAGCAACGGTTCT-3'), for *Chlorella* virus DNA ligase, with primers pCAGRTfw and chlRTPCR-RV1 (5'-CA GCACCTGTGGTGCTTGA-3') and for *E. coli* DNA ligase, with primers pCAGRTfw and LigARTPCR-RV1 (5'-CCTGCACACGTTTGTGAAA-3'). RNA was isolated using RNeasy Mini Kit (Qiagen) and cDNA was generated by the SuperScript III First-strand Synthesis system (Invitrogen).

Genotyping. Genomic DNA was isolated using the Genelute Mammalian Genomic DNA Miniprep Kit (Sigma). Each primer was named for the location on the genomic DNA (for example, Int5-6Fw means that the primer is at the intron between exons 5 and 6). Primer pairs used for genotyping were as follows: Exon 5Fw and Neo2Rv (primer pair 'a' in Fig. 1a), 5'-GGCTTTCACGGTG ATGTGTA-3' and 5'-TCTGGATTTCGACTGTGG-3', using an annealing temperature of 62 °C; Int5-6Fw and Int16-17Rv (primer pair 'b' in Fig. 1a), 5'-CGGGTGTAGGGAGGTCATAA-3', 5'-GAAGGAAGAGGTCTCCAGCA-3', using an annealing temperature of 62 °C; Int10-11Fw and Int11-12Rv (primer pair 'c' in Fig. 1a), 5'-CACTAAACGTGGCAGAGCAA-3', 5'-CCAGCCCCA GACTACAGCTTC-3', using an annealing temperature of 62 °C; Int5-6Fw2

and Int5-6Rv (Supplementary Fig. 1d), 5'-GCCAAGTGTGAATATACAGC-3' and 5'-CAGGGAGCTTGGGACGGATGC-3', using an annealing temperature of 64 °C; Int5-6 and Int16-17 (Supplementary Fig. 1d), 5'-CGGGTGTAGGGA GGTCTATAA-3' and 5'-GAAGGAAGAGGTCTCCAGCA-3', using an annealing temperature of 64 °C.

Microscopy. The subcellular localization of the various GFP fusion constructs was checked by Mitotracker Red CMXRos (Invitrogen) and Hoechst 33342 (Invitrogen) to stain mitochondria and nucleus, respectively. DNA constructs were transiently transfected with Lipofectamine 2000 (Invitrogen). After incubating cells with Opti-MEM (Invitrogen) containing 10 nM Mitotracker Red CMXRos and 2.5 μM Hoechst 33342 for 20 min at 37 °C, cells were monitored with a Zeiss LSM 510 META laser scanning microscope.

qPCR mtDNA repair assay. One million mES cells with the indicated genotypes were plated on 6-cm² plates. After 16 h, cells were cultured with 6.25 ml, 175 μM hydrogen peroxide for 15 min and then cultured with conditioned medium for 1.5 h. mtDNA copy number and mtDNA repair were determined by a long-range qPCR assay¹⁷. Basically, DNA was extracted from pellets of 1×10^6 cells with the DNeasy Blood and Tissue kit (Qiagen) by a QIAcube automated DNA extraction robot (Qiagen). Initial DNA concentration was measured using PicoGreen double-stranded DNA binding agent (Invitrogen) and a DNA standard curve. Total mouse genomic DNA at an approximate final concentration of $4.5 \text{ ng } \mu\text{l}^{-1}$ was then digested with HaeII (New England Biosciences) for 1 h at 37 °C in a reaction mixture containing 1× NEBuffer 4, 1× BSA and 20 U undiluted HaeII enzyme. HaeII linearizes the mouse mtDNA by digesting once (2604) near the D-loop region. Linearization of mitochondrial DNA ensures efficient amplification and allows accurate determination of mtDNA copy number. After digestion, DNA concentration was measured with PicoGreen and an appropriate volume was directly removed from the digest to use for qPCR, with less than 5% variability in DNA concentration between samples.

The qPCR reaction was performed with the GeneAmp XL PCR kit (Applied Biosystems) as follows: 10–15 ng total DNA, in a reaction mix of 50 μl, with 1× buffer, 100 ng μl⁻¹ BSA, dNTPs at 200 μM for each nucleotide, 1.2 and 1.1 mM MgO(Ac)₂ for the long and short fragments, respectively, and 20 pmol for each of two primers. Primer pairs for a 10-kb fragment of mtDNA (long) and for a 117-bp fragment of mtDNA (short) were used for calculating mtDNA damage and mtDNA copy number, respectively. Primer sequences are as described previously¹⁷. DNA polymerase was added at a concentration of 1 U per reaction. A 50% control and a 'no template' blank were used to ensure that the assay was within quantitative range and free of contamination, respectively. PCR products were quantified using fluorescence-blank measurements from the PicoGreen double-stranded DNA binding agent. PCR products from the long fragment were normalized to the short fragment to account for the effect of differing mtDNA copy number on amplification of the long fragment.

SCE. Five million mES cells with the indicated genotypes were plated on 10-cm² plates. After 24 h, cells were cultured with 10 μM bromodeoxyuridine for 20 h (approximately two cell-cycle periods) and pulsed with 0.03 μg ml⁻¹ colcemid for the final 30 min. The cells were collected by centrifugation and exposed to 0.075 mM KCl hypotonic solution for 30 min at 25 °C. The cells were washed twice with the fixative (methanol:acetic acid = 3:1) and suspended in a small volume of the fixative. The cell suspension was dropped onto ice-cold glass slides and air-dried at 60 °C for 2 h. Two days later, slides were incubated with 1 μg ml⁻¹ Hoechst 33258 in Sorensen's phosphate buffer (38 mM KH₂PO₄, 60 mM Na₂HPO₄, pH 7.0) for 10 min, rinsed with 2× SSC buffer (300 mM NaCl, 30 mM Na₂C₂H₃O₇, pH 7.0) and then overlaid with coverslips. Slides were exposed to black light (λ = 352 nm) at a distance of 1 cm for 20 min. After removal of coverslips, the slides were incubated in 2× SSC at 60 °C for 2 h and then stained with 4% (v/v) Giemsa solution in Sorensen's buffer for 10 min, rinsed in water and air-dried. A two-tailed unpaired *t*-test was used to analyse the data.

Drug sensitivity assays. Two thousand mES cells per well were seeded in 24-well plates in duplicate. After 24 h, cells were incubated with various drugs at the indicated concentrations (Fig. 3a) for 24 h in mES cell medium, except hydrogen peroxide which was for 1 h. For irradiation, plates were exposed to the X-ray source from an X-RAD 225C apparatus at a rate of 687 cGy min⁻¹. Six days later, cells were fixed with a solution of 12.5% (v/v) acetic acid, 12.5% (v/v) methanol for 15 min and then stained with 1% (w/v) crystal violet. Afterwards, stained cells were treated with 0.1% SDS in methanol and the absorbance was measured at 595 nm. Each point in the plots was the average of two experiments where each experiment had a duplicate and was a percentage of the absorbance from untreated embryonic stem cells. For Lig3-null cells expressing the Lig3 ΔZNF-GFP, MtLig3-ΔBRCT-GFP-NES and MtLig1 GFP transgenes, two independent null clones were used for each. For the colony formation assays (Fig. 3b), 2×10^3 embryonic stem cells were plated in 10-cm² plates and exposed to ionizing radiation or ultraviolet C radiation. Eight days later, surviving colonies were fixed with methanol and stained with 3% (v/v) Giemsa.

Structural basis of RNA polymerase II backtracking, arrest and reactivation

Alan C. M. Cheung¹ & Patrick Cramer¹

During gene transcription, RNA polymerase (Pol) II moves forwards along DNA and synthesizes messenger RNA. However, at certain DNA sequences, Pol II moves backwards, and such backtracking can arrest transcription. Arrested Pol II is reactivated by transcription factor IIS (TFIIS), which induces RNA cleavage that is required for cell viability¹. Pol II arrest and reactivation are involved in transcription through nucleosomes^{2,3} and in promoter-proximal gene regulation^{4–6}. Here we present X-ray structures at 3.3 Å resolution of an arrested *Saccharomyces cerevisiae* Pol II complex with DNA and RNA, and of a reactivation intermediate that additionally contains TFIIS. In the arrested complex, eight nucleotides of backtracked RNA bind a conserved ‘backtrack site’ in the Pol II pore and funnel, trapping the active centre trigger loop and inhibiting mRNA elongation. In the reactivation intermediate, TFIIS locks the trigger loop away from backtracked RNA, displaces RNA from the backtrack site, and complements the polymerase active site with a basic and two acidic residues that may catalyse proton transfers during RNA cleavage. The active site is demarcated from the backtrack site by a ‘gating tyrosine’ residue that probably delimits backtracking. These results establish the structural basis of Pol II backtracking, arrest and reactivation, and provide a framework for analysing gene regulation during transcription elongation.

Backtracking of bacterial RNA polymerase and the related eukaryotic Pol II to an arrested state is triggered by a weak DNA–RNA hybrid, and dislodges the RNA 3′ end from the active site^{7–9}. A recent study attempted to crystallize an arrested Pol II complex with the use of DNA–RNA scaffolds containing 3′-overhanging RNA¹⁰. This allowed the visualization of one or two backtracked RNA nucleotides, but not further-backtracked RNA¹⁰. We have now resolved further-backtracked RNA nucleotides by crystallization of an arrested complex obtained by Pol II transcription of a 3′-tailed DNA template. Tailed template transcription was previously used for structural studies¹¹, and is prone to arrest¹². We incubated Pol II with a tailed template (Fig. 1a) and different NTP substrates, crystallized the obtained complexes, collected diffraction data, and inspected difference electron density maps after phasing with the free Pol II structure. Only incubation with CTP alone resulted in interpretable difference density for backtracked RNA (Fig. 1b). The register of the nucleic acids was defined by bromine labelling of the –5 template base (Fig. 1a, b) and the structure was refined at 3.3 Å resolution (Methods and Supplementary Table 1).

The structure revealed 13 base pairs (bp) of downstream DNA, a 6-bp hybrid, and 9 nucleotides of single-stranded 3′ RNA that is extruded through the pore into the funnel (Fig. 1). Observation of a 15-nucleotide RNA with 9 unpaired residues is consistent with arrest of tailed template transcription after synthesis of 13–17 nucleotides¹², with a binding site on Pol II for a 9-nucleotide 3′ RNA¹³, and with patterns of RNA cleavage and nuclease protection in arrested complexes^{14,15}. Thus, the template had allowed for CMP mis-incorporation, and the resulting destabilized hybrid drove backtracking to the arrested state.

Comparison with the elongation complex structure reveals that the DNA–RNA hybrid is tilted towards the bridge helix (Fig. 1c). In the

elongation complex, the RNA 3′ end occupies position –1, and the incoming NTP substrate and the templating DNA base occupy position +1, relative to the active site. In the arrested complex, the –1 base pair is tilted by ~25°. Its DNA base resides in the site that is normally occupied by the +1 nucleotide, which instead resides in the downstream cleft, leaving the non-complementary +1 RNA base unpaired. In contrast, the +1 RNA base is paired with the DNA template base in a previously reported complex backtracked by one position¹⁰. Whether hybrid tilting is a cause or consequence of backtracking or whether it results from the shorter hybrid remains to be investigated.

The structure shows that backtracked RNA is extruded into the pore and funnel (Fig. 1d), confirming an early hypothesis¹⁶. The key feature of the complex is the highly defined structure of backtracked RNA and its extended binding site (Fig. 2). Backtracked RNA binds a ‘backtrack site’ along one side of the pore and the mobile trigger loop on the opposite side. The trigger loop has been proposed to control the lateral oscillation of polymerase¹⁷. Its ‘trapped’ conformation observed here is distinct from the closed, open and wedged conformations that occur during forward elongation^{18–21} (Supplementary Fig. 1), and also different from the conformation observed in a complex backtracked by one nucleotide¹⁰. These observations indicate the basis for transcription arrest. When backtracking is not extensive, RNA interactions with the backtrack site are partial and weak, and Pol II can spontaneously move forward and continue elongation. When backtracking is more extensive, backtracked RNA and the trigger loop are trapped, preventing forward elongation and arresting Pol II.

The structure reveals details of backtracked RNA and its interactions with the backtrack site (Fig. 2 and Supplementary Table 2). The first backtracked RNA nucleotide +2 contacts the bridge helix residue T827 and the fork residue E529, and stacks between the +1 base and Rpb2 residue Y769 that we call the ‘gating tyrosine’. The +2 base also contacts the gating tyrosine in a previous structure, although via an edge-to-face interaction¹⁰. Beyond position +2, base stacking discontinues due to steric hindrance by the gating tyrosine. The RNA backbone kinks between nucleotides +2 and +3, and contacts Rpb2 residue R766. The +3 nucleotide binds the trigger loop residues Q1078 and T1080, consistent with crosslinking data²². Beyond position +3, RNA binds exclusively to the Rpb1 funnel domain. A stack of bases from nucleotides +3 to +5 is followed by nucleotide +6, which inserts its base into a pocket that is flanked by R726 and I756 and called here the ‘funnel pocket’. Between positions +6 and +7, the RNA kinks again. The +7 base stacks against residue F755 on one side and against the +8 and +9 bases on the other. Because RNA-binding residues are conserved among eukaryotes (Supplementary Table 2 and Supplementary Fig. 2), backtracked RNA probably binds the same backtrack site in all Pol II enzymes. Modelling purine bases onto the backtracked RNA reveals potential clashes with Pol II (Supplementary Fig. 3). This indicates that the backtrack site preferentially accommodates pyrimidine bases, and provides an explanation for why known DNA arrest sites direct synthesis of pyrimidine-rich RNA²³.

The RNA conformation and Pol II interactions indicate why intrinsic RNA cleavage generally occurs after backtracking by one position, a

¹Gene Center and Department of Biochemistry, Center for Integrated Protein Science Munich, Ludwig-Maximilians-Universität München, Feodor-Lynen-Str. 25, 81377 Munich, Germany.

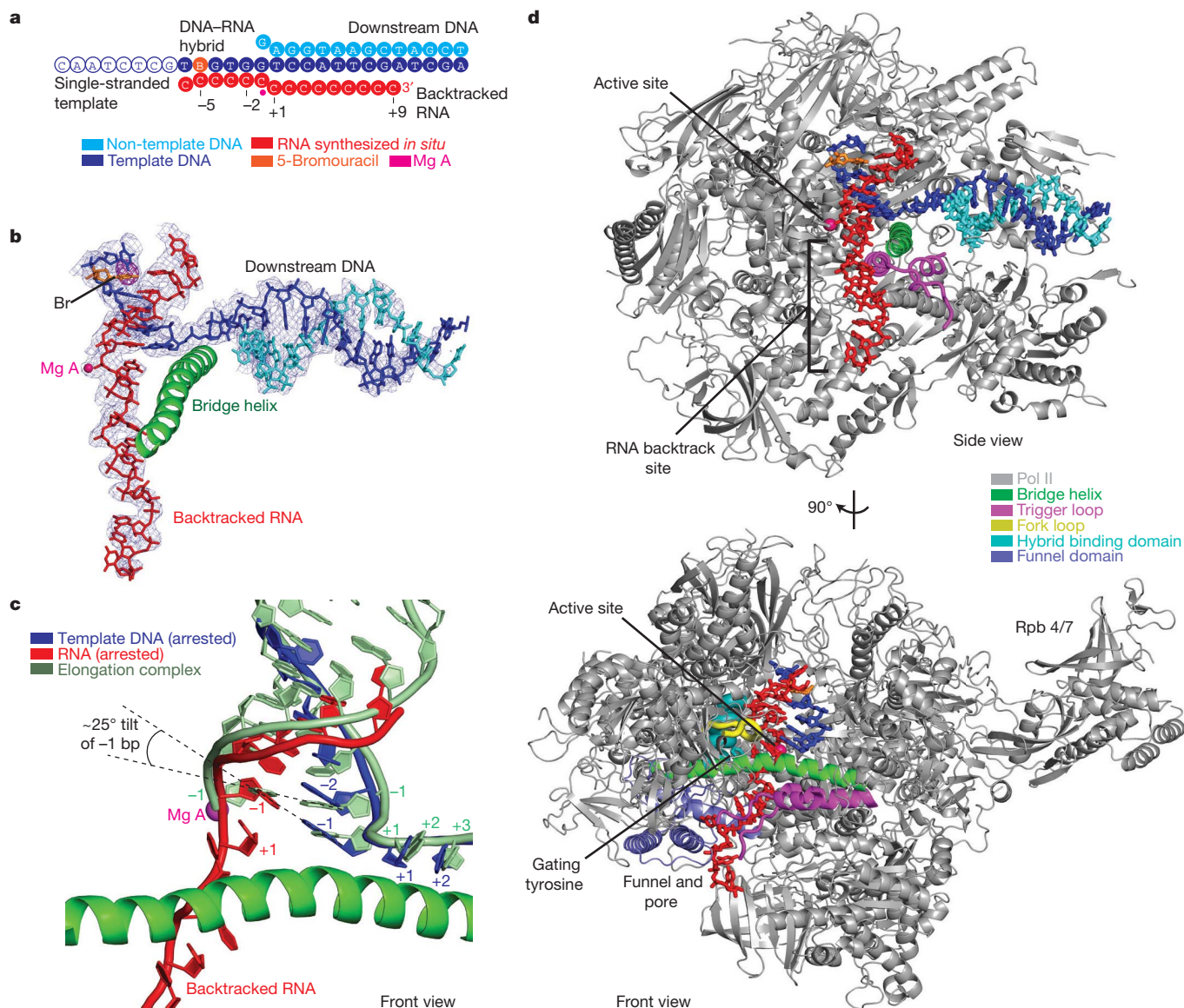


Figure 1 | Structure of arrested Pol II. **a**, Schematic of nucleic acids. Ordered nucleotides are shown with filled circles. The colour code is used throughout. The NTP-binding site corresponds to position +1 and backtracked (downstream) residues are labelled with positive numbers. Some complexes may contain additional disordered nucleotides at the RNA 3' end. **b**, Unbiased difference electron density (blue mesh, contoured at 3.0σ) for nucleic acids,

after phasing with the Pol II structure. A bromine atom is revealed by anomalous difference density (magenta mesh, contoured at 4σ , peak height 7σ) and defined the nucleic acid register. **c**, Front view¹⁶ of a superposition of the tilted DNA-RNA hybrid in the arrested Pol II structure with that in the structure of the elongation complex (light green). **d**, Side and front views of the arrested Pol II complex structure with functional elements highlighted.

phenomenon that underlies mRNA proofreading^{10,24,25}. Backtracking begins with Pol II pausing and fraying of the 3'-terminal RNA nucleotide +1 against the gating tyrosine^{24,26}. The gating tyrosine maintains contact with the 3' nucleotide during the first step of backtracking¹⁰ (Fig. 2a). Backtracking by one position may thus be facilitated, but further backtracking is probably disfavoured because RNA base stacking must be discontinued at the gating tyrosine. Hence, the gating tyrosine may generally delimit the extent of backtracking. If, however, base-stacking interactions and the hybrid are weak, backtracking beyond the gating tyrosine may occur and lead to arrest.

Arrested Pol II is reactivated by TFIIS, which induces cleavage of backtracked RNA. TFIIS binds with its domain II near the rim of the Pol II funnel, and extends into the Pol II pore with its domain III, which reaches the active site with a β -stranded hairpin^{19,27}. Superposition of the arrested complex with the Pol II-TFIIS complex²⁷ revealed that the backtracked RNA overlaps with TFIIS domain III. This indicated that backtracked RNA prevents TFIIS from invading the pore, posing the

question of how reactivation occurs. To investigate this, we soaked arrested complex crystals with a TFIIS variant that carries two point mutations in functionally essential acidic hairpin residues (D290A/E291A), and solved the structure of the resulting reactivation intermediate at 3.3 Å resolution (Methods). As observed before¹⁹, TFIIS changed the Pol II conformation, realigned RNA in the hybrid, inserted its domain III into the pore, and reached the active site with its hairpin.

In the structure of the reactivation intermediate, TFIIS binding moved the trigger loop by up to 5 Å from the trapped to the 'locked' position (Fig. 3a) observed previously²⁷. TFIIS binding also induces a rotation of the gating tyrosine side chain, which prevents its stacking with backtracked RNA. Backtracked RNA was displaced from the backtrack site into the part of the pore that remains accessible after binding of TFIIS domain III ('restricted pore', Fig. 3b). The electron density for displaced RNA was discontinuous and only allowed for backbone modelling (Fig. 3b). The discontinuous density apparently

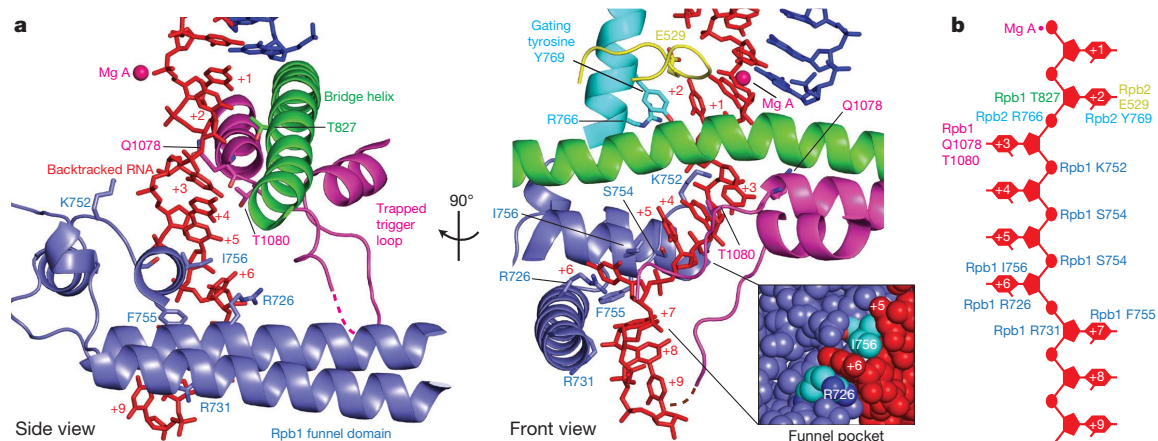


Figure 2 | Backtracked RNA in the backtrack site. **a**, Side and front views of backtracked RNA. RNA-binding Pol II elements in the pore and funnel are depicted, and contact residues in the backtrack site are labelled. **b**, Schematic of Pol II interactions with backtracked RNA.

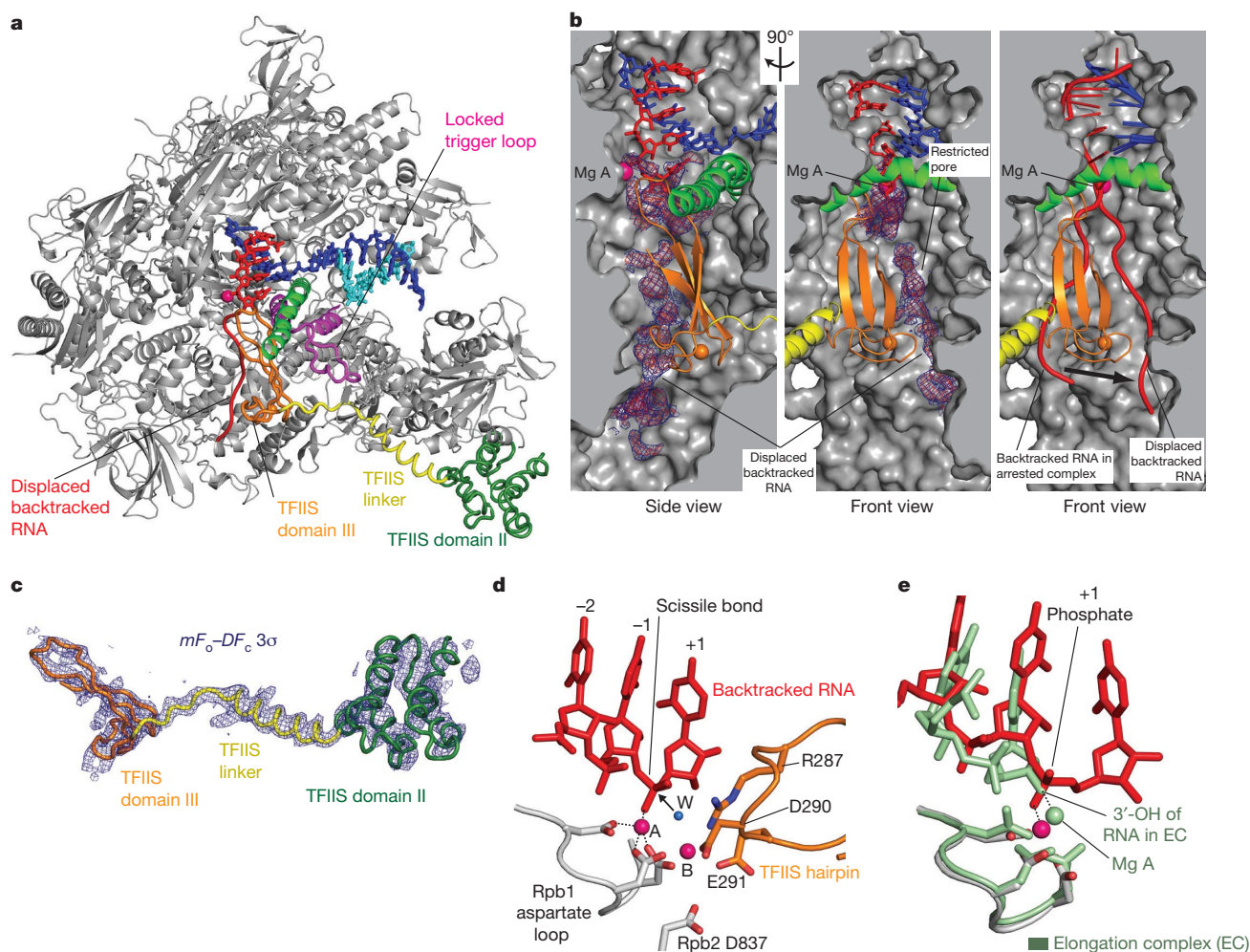


Figure 3 | Structure of reactivation intermediate. **a**, Side view. TFIIS domains II and III are in green and orange, respectively, and the domain II–III linker is in yellow. **b**, Displacement of backtracked RNA from the backtrack site into the restricted pore. Difference electron density for displaced backtracked RNA is contoured at 2σ (blue) and 2.5σ (red). On the right, backtracked RNA from the arrested complex is modelled into the structure, revealing a clash with TFIIS domain III. **c**, Unbiased difference electron density for TFIIS after phasing with the Pol II structure contoured at 3σ , and sigmaA-weighted by

coefficients m and D . **d**, Model for active site geometry during TFIIS-induced RNA cleavage. Metal A is coordinated by the Pol II aspartate loop and the RNA +1 phosphate. Metal B, the nucleophilic water molecule (W, blue sphere), and side chains of TFIIS hairpin residues D290 and E291 (orange) are placed onto the crystal structure without clashes. This supports a catalytically competent arrangement for an S_N2 mechanism. The arrow indicates the direction of the in-line nucleophilic attack. **e**, Superposition of the view in **d** with the active site conformation of the Pol II elongation complex¹⁹.

reflects rotational mobility rather than partial release of backtracked RNA, as RNA was not cleaved in a reconstituted backtracked complex by the inactive TFIIS variant (not shown). These observations indicate that TFIIS weakens the Pol II grip on backtracked RNA, and displaces and mobilizes RNA by competitive binding to the backtrack site.

Pol II reactivation by TFIIS-dependent cleavage apparently involves two metal ions and a nucleophilic water molecule^{27–29}. Metal A is persistently bound to the active site aspartate loop¹⁶, whereas metal B may be recruited by TFIIS^{27,29}. The reactivation intermediate structure provides new insights into TFIIS-induced cleavage, as it is at higher resolution than previous work^{10,19}, and because it reveals the RNA +1 nucleotide, and thus the scissile phosphodiester bond between nucleotide –1 and +1 (Fig. 3d). First, metal A binds the +1 RNA phosphate to align the scissile bond, in contrast to its binding of the RNA 3' hydroxyl during nucleotide addition (Fig. 3e). Second, the TFIIS hairpin residue R287 reaches the catalytic site and could stabilize the negatively charged transition state, explaining its role in catalysis³⁰. Third, modelling the side chains of D290 and E291 onto the structure

indicates that they can coordinate metal B, together with the Rpb2 residue D837 (Fig. 3d). The invariant charged TFIIS residues R287, D290 and E291 may be required for two catalytic proton transfers, proton subtraction from the nucleophilic water, and proton donation to the product RNA 3' terminus.

The presented structural snapshots of transcription intermediates reveal the mechanisms of Pol II backtracking, arrest and reactivation (Fig. 4). When Pol II encounters a DNA sequence that impairs elongation, it pauses and backtracks by one position. Because further backtracking is hindered by the gating tyrosine, polymerase-intrinsic cleavage of a dinucleotide from the RNA 3' end can occur, and elongation continues. However, at an arrest site the hybrid is weak and the RNA can backtrack beyond the gating tyrosine. Extensive backtracking traps RNA and the trigger loop in the pore, inhibiting elongation and arresting Pol II. TFIIS reactivates arrested Pol II by locking the trigger loop away from RNA, displacing and mobilizing backtracked RNA in the pore, and complementing the active site with a basic and two acidic side chains. This induces cleavage and release of backtracked RNA, and creates a new RNA 3' end at the active site that allows transcription to resume.

METHODS SUMMARY

Saccharomyces cerevisiae 12-subunit Pol II was prepared as described²⁴. Purified Pol II (3.5 mg ml^{–1}) was mixed with a twofold molar excess of tailed template (Fig. 1a) prepared as described¹⁹, 8 mM magnesium chloride and 2 mM CTP, and incubated for 2 h at 20 °C before crystallization by vapour diffusion with 6% PEG 6000, 200 mM ammonium acetate, 300 mM sodium acetate, 50 mM HEPES pH 7.0 and 5 mM TCEP as reservoir solution. Crystals were grown for 5–10 days, cryo-protected in mother solution supplemented with 22% glycerol and containing 2 mM CTP, 8 mM magnesium chloride, and 4 μM tailed template, followed by overnight incubation at 8 °C before harvesting and freezing in liquid nitrogen. Arrested Pol II–TFIIS complex crystals were prepared by adding the inactive TFIIS variant D290A/E291A²⁷ to the cryo-protectant at 1 mg ml^{–1} and incubating arrested Pol II crystals overnight at 8 °C. Diffraction data at 3.3 Å were collected at 100 K at beamline X06SA of the Swiss Light Source and structures were solved with molecular replacement using the model of 12-subunit Pol II (1WCM). Data were collected at 13.494 keV, the K-absorption peak of bromine.

Received 13 October; accepted 23 December 2010.

Published online 23 February 2011.

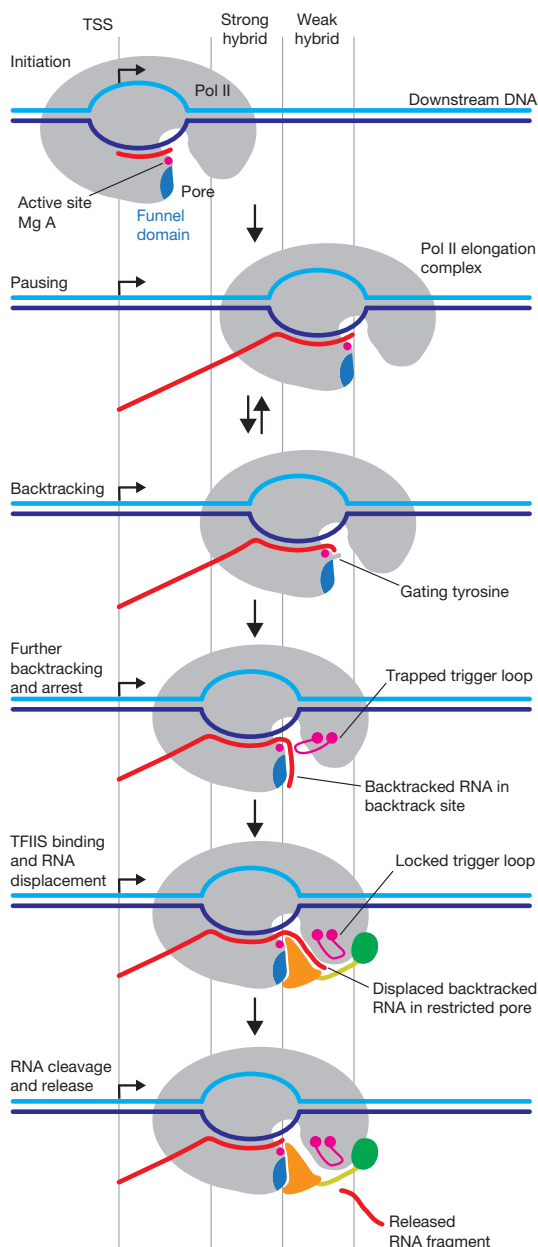


Figure 4 | Mechanism of Pol II backtracking, arrest and reactivation. Schematic of Pol II functional states. For details, compare text.

1. Sigurdsson, S., Dirac-Svejstrup, A. B. & Svejstrup, J. Q. Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. *Mol. Cell* **38**, 202–210 (2010).
2. Kim, J., Guermah, M. & Roeder, R. G. The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS. *Cell* **140**, 491–503 (2010).
3. Kireeva, M. L. *et al.* Nature of the nucleosomal barrier to RNA polymerase II. *Mol. Cell* **18**, 97–108 (2005).
4. Adelman, K. *et al.* Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. *Mol. Cell* **17**, 103–112 (2005).
5. Nechaev, S. *et al.* Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* **327**, 335–338 (2010).
6. Palangat, M., Renner, D. B., Price, D. H. & Landick, R. A negative elongation factor for human RNA polymerase II inhibits the anti-arrest transcript-cleavage factor TFIIS. *Proc. Natl Acad. Sci. USA* **102**, 15036–15041 (2005).
7. Komissarova, N. & Kashlev, M. RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA. *J. Biol. Chem.* **272**, 15329–15338 (1997).
8. Nudler, E., Mustaev, A., Lukhtanov, E. & Goldfarb, A. The RNA–DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* **89**, 33–41 (1997).
9. Palangat, M. & Landick, R. Roles of RNA:DNA hybrid stability, RNA structure, and active site conformation in pausing by human RNA polymerase II. *J. Mol. Biol.* **311**, 265–282 (2001).
10. Wang, D. *et al.* Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. *Science* **324**, 1203–1206 (2009).
11. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**, 1876–1882 (2001).
12. Sluder, A. E., Price, D. H. & Greenleaf, A. L. Elongation by *Drosophila* RNA Polymerase-II. Transcription of 3'-extended DNA templates. *J. Biol. Chem.* **263**, 9917–9925 (1988).
13. Johnson, T. L. & Chamberlin, M. J. Complexes of yeast RNA polymerase II and RNA are substrates for TFIIS-induced RNA cleavage. *Cell* **77**, 217–224 (1994).

14. Gu, W., Powell, W., Mote, J. J. & Reines, D. Nascent RNA cleavage by arrested RNA polymerase II does not require upstream translocation of the elongation complex on DNA. *J. Biol. Chem.* **268**, 25604–25616 (1993).
15. Izban, M. G. & Luse, D. S. The increment of SII-facilitated transcript cleavage varies dramatically between elongation competent and incompetent RNA polymerase II ternary complexes. *J. Biol. Chem.* **268**, 12874–12885 (1993).
16. Cramer, P. *et al.* Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**, 640–649 (2000).
17. Bar-Nahum, G. *et al.* A ratchet mechanism of transcription elongation and its control. *Cell* **120**, 183–193 (2005).
18. Brueckner, F. & Cramer, P. Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nature Struct. Mol. Biol.* **15**, 811–818 (2008).
19. Kettenberger, H., Armache, K.-J. & Cramer, P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIIS. *Mol. Cell* **16**, 955–965 (2004).
20. Vassylyev, D. G. *et al.* Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**, 163–168 (2007).
21. Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D. & Kornberg, R. D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941–954 (2006).
22. Markovtsov, V., Mustaev, A. & Goldfarb, A. Protein-RNA interactions in the active center of transcription elongation complex. *Proc. Natl Acad. Sci. USA* **93**, 3221–3226 (1996).
23. Hawryluk, P. J., Ujvari, A. & Luse, D. S. Characterization of a novel RNA polymerase II arrest site which lacks a weak 3' RNA-DNA hybrid. *Nucleic Acids Res.* **32**, 1904–1916 (2004).
24. Sydow, J. F. *et al.* Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol. Cell* **34**, 710–721 (2009).
25. Zenkin, N., Yuzenkova, Y. & Severinov, K. Transcript-assisted transcriptional proofreading. *Science* **313**, 518–520 (2006).
26. Touloukhanov, I., Zhang, J. W., Palangat, M. & Landick, R. A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol. Cell* **27**, 406–419 (2007).
27. Kettenberger, H., Armache, K.-J. & Cramer, P. Architecture of the RNA polymerase II-TFIIIS complex and implications for mRNA cleavage. *Cell* **114**, 347–357 (2003).
28. Sosunov, V. *et al.* Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J.* **22**, 2234–2244 (2003).
29. Weillbaecher, R. G., Awrey, D. E., Edwards, A. M. & Kane, C. M. Intrinsic transcript cleavage in yeast RNA polymerase II elongation complexes. *J. Biol. Chem.* **278**, 24189–24199 (2003).
30. Awrey, D. E. *et al.* Yeast transcript elongation factor TFIIIS, structure and function. II: RNA polymerase binding, transcript cleavage, and read-through. *J. Biol. Chem.* **273**, 22595–22605 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank F. Brueckner, G. Damsma, K. Kinkelin, D. Kostrewa, L. Larivière, E. Lehmann, F. Martinez, S. Sainsbury and J. Sydow. Part of this work was performed at the Swiss Light Source at the Paul Scherrer Institut, Villigen, Switzerland. P.C. was supported by the Deutsche Forschungsgemeinschaft, SFB646, TR5, FOR1068, NIM, Bioimaging Network BIN, and the Jung-Stiftung.

Author Contributions A.C.M.C. carried out experiments. P.C. supervised the project. A.C.M.C. and P.C. prepared the manuscript.

Author Information Coordinates and structure factors of the arrested Pol II elongation complex and the arrested Pol II reactivation intermediate have been deposited with the Protein Data Bank under accession numbers 3PO2 and 3PO3, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.C. (cramer@lmb.uni-muenchen.de).

CORRIGENDUM

doi:10.1038/nature09877

Evolution of human *BCR-ABL1* lymphoblastic leukaemia-initiating cells

Faiyaz Notta, Charles G. Mullighan, Jean C. Y. Wang,
Armando Poepl, Sergei Doulatov, Letha A. Phillips, Jing Ma,
Mark D. Minden, James R. Downing & John E. Dick

Nature **469**, 362–367 (2011)

On page 2 of the PDF and print version of this Article, column one, under the heading 'Genetic basis of functional heterogeneity', line 9 should refer to Fig. 1e instead of Fig. 2e. Similarly, in column two of the same section, lines 7 and 9 should ref to Fig. 1f and 1g, respectively, instead of Fig. 2f and 2g. These have now been corrected in the HTML version of the manuscript.

CORRECTIONS & AMENDMENTS

CORRIGENDUM

doi:10.1038/nature09858

Reliability of flipper-banded penguins as indicators of climate change

Claire Saraux, Céline Le Bohec, Joël M. Durant, Vincent A. Viblanc, Michel Gauthier-Clerc, David Beaune, Young-Hyang Park, Nigel G. Yoccoz, Nils C. Stenseth & Yvon Le Maho

Nature **469**, 203–206 (2011)

In this Letter, the statement in the abstract and main text that the banded birds had a survival rate “16% lower than non-banded birds” should in fact say that they had a survival rate “16 percentage points lower than non-banded birds”. Also, the abstract should state that banded birds had 41% fewer chicks, rather than 39%.

CORRIGENDUM

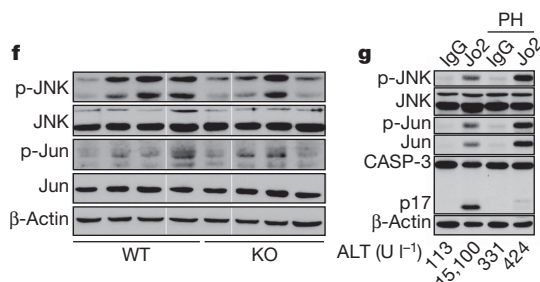
doi:10.1038/nature09897

CD95 promotes tumour growth

Lina Chen, Sun-Mi Park, Alexei V. Tumanov, Annika Hau, Kenjiro Sawada, Christine Feig, Jerrold R. Turner, Yang-Xin Fu, Iris L. Romero, Ernst Lengyel & Marcus E. Peter

Nature **465**, 492–496 (2010)

In this Letter, an experimental error affected the western blot analyses of mouse liver extracts shown in Fig. 4f and g. The secondary antibody cross-reacted with endogenous mouse IgG in the tissue lysates, resulting in incorrect bands. The experiments were repeated using different primary antibodies and a secondary antibody that showed no such cross-reactivity. Although there was a trend, the reduction of p-JNK and p-Jun in the livers of CD95-deficient mice in Fig. 4f was not statistically significant. In addition, although the increase in phosphorylated JNK and Jun in Jo2-injected mice was reproducible in Fig. 4g, Jun protein was also increased. Corrected versions of Fig. 4f and g are shown below. The Supplementary Methods have been updated to reflect the change in the use of antibodies, and Supplementary Fig. 8b has also been corrected. The corrected figures do not alter the overall conclusions of this Letter, and all other data still stand.



THE CAMBRIAN

It doesn't add up.

BY GEORGE ZEBROWSKI

We were sure that Hinkle had to be deluded when he first told us that he had been talking to horseshoe crabs.

"They are intelligent," he insisted, "a variation of the oldest on our planet. They think about a lot of things."

He showed me equations.

"You did these yourselves," I said. "Come on, admit it."

Physics was my game, but his maths gave me a strange feeling, familiar and unfamiliar at the same time.

"Where did you get these?" I asked.

"From the Cambrians."

There weren't any equations like this that I had ever seen.

"If it's a delusion," I told Hinkle, "then you need to keep at it," half joking.

"What do you mean?"

"You deserve the credit, of course," I said.

"But I have no idea of what they mean, or what they're about," Hinkle said.

"You must know," I said, "or you couldn't have written them."

"I only took them down," he insisted.

"Why are you lying?"

"You're my friend. Why should I lie to you?"

"Who are you protecting?"

"No one. The crabs talk to me. Here, look."

The equations all led to one unified statement about the Universe, but I noted that there seemed to be an incompleteness about the final statement, and wondered what crank book he had found them in.

"What about that?" I asked him.

"Search me," he said. "I'm just a conveyor."

"So you say," I said, "but this claim is so comprehensive and suggestive... yet incomplete."

"Remember, I am a palaeobiologist first."

I nodded and said: "Stop kidding me."

"Well, I can tell you," Hinkle said as he took a deep breath, "that I suspect it's supposed to be incomplete, but that it says all that can be said about our Universe. Who was it that said that science attempts to say something about our Universe, but will never say everything that can be said?"

"Bohr, maybe Feynman," I said. "Doesn't matter. But do you have

any idea of the assertions you've written down?"

"You tell me," Hinkle said.

"Either physics can be completed, or it can never have an end, and this final equation of yours asserts a *perpetuum mobile*."

"It's news to me. Enzymes interest me, especially as they affect intelligence."

I gave him a hostile glare, the first in our long friendship, and it seemed to shake him.

"Who's behind all this? I demanded.

"What do you mean?" Hinkle asked, looking like a deer in headlights.

"I get it," I said. "More Gödel incompleteness! Dyson's claim that the theorem is a cathedral of implications, that he would feel sorry for physics if it could be ended, a cultural disaster, leaving us with technology and a manipulation of the Universe through instrumentalities, but nothing more to learn, ever. I'm tired of Gödel and his nothing proof."

"I wouldn't know," Hinkle said, "but I never thought any such thing. These critters are the oldest thinkers on Earth, and they don't wish us any harm."

"But you are doing harm, with these pseudo equations!" I cried. "Don't you know that you can't just claim such things?"

"Then leave me out of it," Hinkle said, "as all this doesn't mean anything, as you say."

"Oh, you're sly," I said. "You know damn well that we have enough dummies who may take it seriously. You're sly, all right."

"But if they're nothing at all, then you have nothing to fear."

I looked at my friend and said: "I'll tell you the truth. You fell in love with these crabs, and that seemed to help you come up with all this. That's all there is to it. You can't take it seriously. You obviously know more maths than I thought."

"But I don't," Hinkle said. "One way or another, I don't know."

I looked at him and said: "Show me how it happens, when you write down what they... give you?"

"Sure, look here, I feel one coming in."

I watched him scribble on the back of an envelope, as fluidly as any good mathematician.

I looked at the result and said: "It's a literate mathematical statement, but it has to be a guess."

More equations came, after which he seemed worn out. I packed them in my briefcase and left him sitting there with his eyes closed like an exhausted lover.

Physics continued completing itself. As usual, open ends had to be irrational avenues; infinities had to be renormalized, limited, outlawed.

Yet in his own way, Hinkle became the greatest physicist of his generation, despite a vast suspicion that he had unknown collaborators feeding him variations on anomalous Higgs boson regimes, which he described and applied without the Large Hadron Collider's data.

Patents are granted to him, and his ideas lead to inventions that work but are not explained. "Just luck," many say. "Quantum engineering works, it predicts and needs no explanations, only descriptions."

The last time I went to see him he told me that human and crab brains might have merged. I told him that was an insane idea. "Well, of course, maybe it didn't happen," he said, "but it should have."

He's a fraud of some kind, but we don't know how he does it.

He's a sight on campus, walking a lot, even at his age. People say he looks younger than some years ago. He strides around in a long coat, hunched forward a bit as if there was something on his back. ■

George Zebrowski is an award-winning novelist, story writer, essayist, editor and lecturer. His latest novel is *Empties* (Golden Gryphon Press). *Sentinels In Honor of Arthur C. Clarke*, with Gregory Benford (Hadley Rille Books), is his new anthology.

NATURE.COM
Follow Futures on
Facebook at:
[go.nature.com/mtodm](https://www.facebook.com/go.nature.com/mtodm)



Caspase signalling controls microglia activation and neurotoxicity

Miguel A. Burguillos^{1,2,3}, Tomas Deierborg³, Edel Kavanagh¹, Annette Persson⁴, Nabil Hajji^{1†}, Albert Garcia-Quintanilla⁵, Josefina Cano⁵, Patrik Brundin³, Elisabet Englund⁴, Jose L. Venero² & Bertrand Joseph¹

Activation of microglia and inflammation-mediated neurotoxicity are suggested to play a decisive role in the pathogenesis of several neurodegenerative disorders. Activated microglia release pro-inflammatory factors that may be neurotoxic. Here we show that the orderly activation of caspase-8 and caspase-3/7, known executioners of apoptotic cell death, regulate microglia activation through a protein kinase C (PKC)- δ -dependent pathway. We find that stimulation of microglia with various inflammogens activates caspase-8 and caspase-3/7 in microglia without triggering cell death *in vitro* and *in vivo*. Knockdown or chemical inhibition of each of these caspases hindered microglia activation and consequently reduced neurotoxicity. We observe that these caspases are activated in microglia in the ventral mesencephalon of Parkinson's disease (PD) and the frontal cortex of individuals with Alzheimer's disease (AD). Taken together, we show that caspase-8 and caspase-3/7 are involved in regulating microglia activation. We conclude that inhibition of these caspases could be neuroprotective by targeting the microglia rather than the neurons themselves.

Numerous *in vivo* clinical imaging and neuropathology studies suggest that activated microglia, the resident immune cells of the central nervous system, play prominent roles in the pathogenesis of neurodegenerative disorders, including PD, multiple sclerosis and AD^{1,2}. Microglia are necessary for normal brain function; however, uncontrolled and over-activated microglia can trigger neurotoxicity. They are a prominent source of pro-inflammatory factors and oxidative stress such as tumour-necrosis factor (TNF)- α , nitric oxide and interleukin (IL)-1 β , which are neurotoxic^{2,3}.

Toll-like receptors (TLRs) are a family of pattern-recognition receptors in the innate immune system. Exogenous and endogenous TLR ligands activate microglia⁴. Intracerebral delivery of lipopolysaccharide (LPS), the major component of Gram-negative bacterial walls and a ligand for TLR4, leads *in vivo* to microglia activation and neuronal injury, and is used as model for brain inflammation^{4,5}. Synergistic effects between interferon- γ (IFN- γ) and several TLR ligands (including TLR4) have been suggested, suggesting crosstalk between these pro-inflammatory receptor signalling pathways⁶. Furthermore, IFN- γ receptor-deficient mice are less susceptible to LPS-induced endotoxic shock than control mice⁷. Finally, TLR4 has been implicated in AD pathophysiology in several contexts. Thus, the upregulation of cytokines is TLR4 dependent in an AD mouse model⁸; certain TLR4 single nucleotide polymorphisms are associated with increased risk for AD⁹; the levels of TLR4 messenger RNA (mRNA) are upregulated in APP transgenic mice¹⁰; and increased TLR4 expression is associated with amyloid plaque deposition in AD brain tissue¹⁰.

Caspases, a family of cysteinyl-aspartate-specific proteases, are executioners of apoptotic cell death and their activation is considered a commitment to cell death^{11,12}. Certain caspases, for example caspase-1, also play a pivotal role in immune-mediated inflammation. In this situation, caspase activation is associated with the maturation of pro-inflammatory cytokines, such as IL-1 β , IL-18, IL-33, and not with apoptosis¹³. Inhibition of caspase activation protects against

neuronal loss in several animal models of brain diseases involving activated microglia, including hypoxic ischaemia/stroke, acute bacterial meningitis, brain trauma and 6-hydroxydopamine and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-lesioned parkinsonism models^{2,14–17}. Currently, it is unclear whether inhibition of caspase activation specifically in microglia contributes to the neuroprotective effects of caspase inhibitors. We have now discovered that microglial activation in cell and animal models of inflammation involves caspases and that inhibition of the cascade in microglia prevents neurodegeneration. Furthermore, we demonstrate that caspase activation occurs in microglia in the brains of individuals with PD and AD, and thereby we validate the observations we made in relevant cell and animal models.

Results

Caspase-3/7 control microglia activation

We stimulated BV2 cells with LPS to investigate the molecular pathways involved in microglia activation (Fig. 1). LPS treatment induced caspase-3 cleavage and D(OMe)E(OMe)VD(OMe)-ase (DEVD-ase) activity after only 4 h in BV2 microglia cells in a time- and dose-dependent manner (Fig. 1a and Supplementary Figs 2a, b, e, f and 5d, e). Increased DEVD-ase activity was also observed upon treatment with other pro-inflammogens such as lipoteichoic acid (LTA, TLR2 agonist), PamC3sk4 (synthetic lipopeptide TLR1/2 agonist) and interferon-gamma (IFN- γ) (Supplementary Fig. 3a). This activity primarily reflects caspase-3/7 activities. Both caspase 3 and 7 are known as major apoptosis executioners. Despite the increase in DEVD-ase activity, we did not observe major microglia cell death within the first 24 h after initiating LPS treatment (Fig. 1c and Supplementary Figs 2c and 4a–c, e, f) nor with LTA, PamC3sk4 and IFN- γ treatments (Supplementary Fig. 3b, c). Thus, the LPS-induced increase in caspase 3 and 7 activity did not result in major cell death and the little cell death occurring at 48 h was not prevented using the caspase-3/7 inhibitor Z-D(OMe)E(OMe)VD(OMe)-FMK

¹Department of Oncology-Pathology, Cancer Centrum Karolinska, Karolinska Institutet, 171 76, Stockholm, Sweden. ²Departamento de Bioquímica y Biología Molecular, Facultad de Farmacia, Universidad de Sevilla, and Instituto de Biomedicina de Sevilla, 41012 Sevilla, Spain. ³Neuronal Survival Unit, Wallenberg Neuroscience Center, Department of Experimental Medical Science, 221 84 Lund, Sweden. ⁴Department of Pathology, Division of Neuropathology, Lund University Hospital, 221 85 Lund, Sweden. ⁵Servicio de Biología, Centro de Investigación, Tecnología e Innovación, Universidad de Sevilla (CITIUS), 41012 Sevilla, Spain. [†]Present address: Department of Experimental Medicine and Toxicology, Division of Investigative Science, Imperial College London, UK.

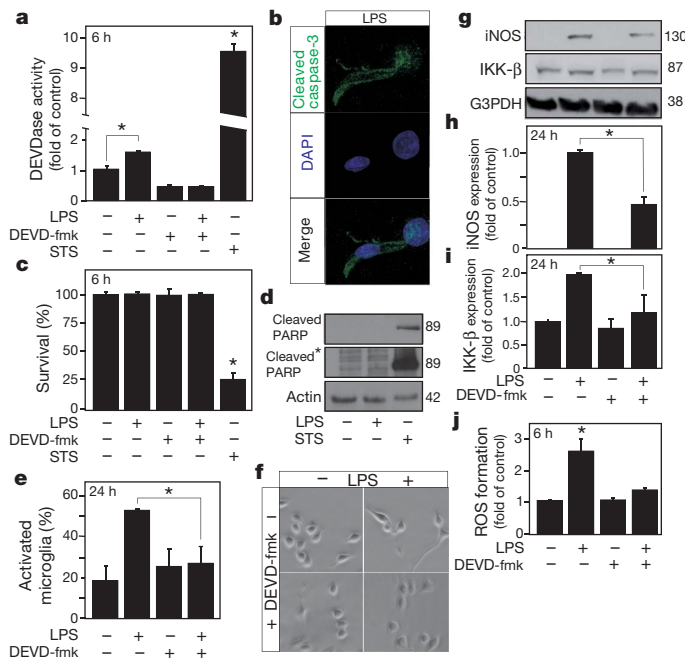


Figure 1 | LPS-induced DEVDase activity regulates microglia activation but not cell death. LPS treatment induces DEVDase activity (a) and processing of caspase-3 (b), which are not associated with cell death, as illustrated by cell survival quantification (c), and absence of PARP cleavage (d) in cultured BV2 microglia cells. Caspase-3/7 inhibition by DEVD-fmk (a) prevents LPS-induced morphological microglia activation (e, f) and induction of typical inflammation-related molecules like iNOS (g, h) and IKK- β (g, i) and ROS formation (j). STS is used as an apoptosis inducer. Data are expressed as mean \pm s.e.m. ($n = 3$). * $P < 0.05$.

(DEVD-fmk) (Supplementary Figs 2d and 4d, g). Compared with LPS treatment, exposure of BV2 microglia cells to a death stimulus such as staurosporine (STS) led to a significantly greater caspase-3 cleavage and induction of DEVD-ase activity (Fig. 1a and Supplementary Fig. 5d, e). After LPS treatment, we found cleaved caspase-3 to be located primarily close to the plasma membrane and not present in the nucleus (Fig. 1b, Supplementary Fig. 5a–d and Supplementary Movies 1 and 2). Furthermore, we did not observe cleavage of the caspase-3/7 nuclear substrate poly(ADP-ribose) polymerase (PARP-1) in response to LPS (Fig. 1d) even after long exposure of the membrane (asterisk in Fig. 1d). In addition, whereas exposure of BV2 cells to STS promoted Bid processing and loss of mitochondrial transmembrane potential, these two events associated with the mitochondrial cell death pathway were found to be unaffected upon LPS treatment (Supplementary Fig. 10a–c). When we inhibited DEVD-ase activity in BV2 cells by exposing them to the cell-permeable and irreversible caspase inhibitor DEVD-fmk, LPS treatment failed to activate the microglia. Thus, the cells did not exhibit morphological changes associated with microglia activation (Fig. 1e–f) and did not show features of activated microglia, such as I κ B kinase complex β (IKK- β), inducible nitric oxide synthases (iNOS) and reactive oxygen species (ROS) formation (Fig. 1g–j). Co-treatment with DEVD-fmk also prevented the LTA, PamC3sk4 and IFN- γ -induced iNOS expression and ROS formation (Supplementary Fig. 6a–e).

Caspase-3 and -7 account for cellular DEVD-ase activity. We therefore decided to assess their respective roles in microglia activation by selectively knocking down endogenous caspase-3 (Supplementary Fig. 7a, b) or caspase-7 (Supplementary Fig. 7c, d) using a pool of small interfering RNAs (siRNAs). First, we confirmed that the silencing of these proteases effectively decreased DEVD-ase activity (Supplementary Fig. 7e). Then, when we transfected BV2 microglia cells with siRNA targeting specifically either one of the two caspases, LPS-treatment did not induce iNOS, IKK- β expression, ROS formation and production of certain cytokine production (IL-1 β , TNF- α and

murine keratinocyte chemoattractant) as effectively (Fig. 2a, b and Supplementary Fig. 7f–i). We found that silencing both caspases simultaneously reduced IKK- β expression even further (Supplementary Fig. 7f). This indicates that DEVD-ase activity per se regulates microglia activation (Fig. 2 and Supplementary Fig. 7). Similar responses to caspase-3 and -7 knockdowns were observed with LTA, PamC3sk4 and IFN- γ treatments (Supplementary Fig. 8a–i). Activation and nuclear translocation of nuclear factor κ B (NF- κ B) is a key step in LPS-induced microglia activation¹⁸. NF- κ B is sequestered in the cytoplasm by the I κ B family of inhibitory proteins that mask the nuclear localization signal of NF- κ B. IKK- β can phosphorylate I κ B, and thereby target it for degradation through the ubiquitin proteasome pathway. As a consequence, functional NF- κ B molecules then become free to enter the nucleus. We detected less nuclear NF- κ B p65 subunits in cells subjected to knockdown of caspase-3 or caspase-7 before LPS treatment (Fig. 2c, d), indicating that reduced caspase activation led to less nuclear translocation of NF- κ B. Finally, using microglia cells co-cultured with dopaminergic neurons, we examined whether the inhibition of the IKK/NF- κ B pathways by selective knockdown of caspase-3 or/and caspase-7 was associated with loss of microglia neurotoxicity. In agreement with earlier studies¹⁹, LPS treatment activated microglia and caused dopaminergic neurons to die. Importantly, we found that reducing the LPS-induced microglia activation by caspase knockdown meant that fewer co-cultured dopaminergic neurons died (Fig. 2e, f). Also, we checked the level of several cytokines in primary microglial cell cultures at 12 and 24 h, and observed a decrease of IL-1 β and IL-5 with both DEVD-fmk and IETD-fmk, and of IL-2 and IFN- γ at 12 h when we used IETD-fmk (Fig. 2g).

Caspase-8 directs caspase-3/7 activation

We then examined how caspase-3/7 are activated in LPS-treated microglia. LPS treatment has been reported to promote caspase-1 activation²⁰. Caspase 1 is a key component of the inflammasome required for the processing and maturation of pro-inflammatory cytokines. It plays a pivotal role during LPS-induced inflammation. Therefore caspase-1 null mice and mice expressing a dominant-negative mutant caspase-1 gene exhibit reduced LPS-induced inflammation^{21,22}. Consequently, we examined if caspase-1 acts upstream of caspase-3/7 in LPS-induced microglia activation. We found that the caspase-1 inhibitor YVAD-fmk did not reduce LPS-induced activation of microglia, as assessed by DEVD-ase activity (Fig. 3a).

Caspase-8 is believed to be at the apex of the death receptor-mediated apoptosis pathway and can activate caspase-3/7 (refs 23–25). We found that LPS induced caspase-8 activity (IETD-ase) within 6 h of being added to BV2 microglia cultures (Fig. 3b, c). Consistent with this result, we observed that the caspase-8 inhibitor IETD-fmk or caspase-8 knockdown using siRNA prevented LPS-induced DEVD-ase activity (Fig. 3a, d). The IETD-fmk treatment or caspase-8 silencing only prevented the LPS-induced increase in DEVD-ase activity, whereas the caspase inhibitor DEVD-fmk reduced the DEVD-ase activity even further (Fig. 3a, c). Knockdown or chemical inhibition of caspase-8 were also associated with a reduction of iNOS expression and ROS formation upon treatment with all tested pro-inflammogens (that is, LPS, LTA, PamC3sk4 and IFN- γ) (Fig. 3e, f and Supplementary Figs 7i, 8a–i and 9a, b). Other known potential caspase-8 substrates, namely Bid, HDAC7 and RIP1, were found not to be processed upon LPS treatment (Supplementary Fig. 10 b–f). Taken together, this indicates that the LPS-induced increase in DEVD-ase activity is dependent on caspase-8 activation. Furthermore, we found that the caspase-8- and caspase-1-initiated pathways had additive effects in the regulation of LPS-induced iNOS expression (Supplementary Fig. 9c, d). This suggests that both of these caspase-regulated pathways contribute to LPS-induced inflammation. We also examined if the observed DEVD-ase activity and that of caspase 6, the remaining executioner caspase, were correlated. However, we were unable to

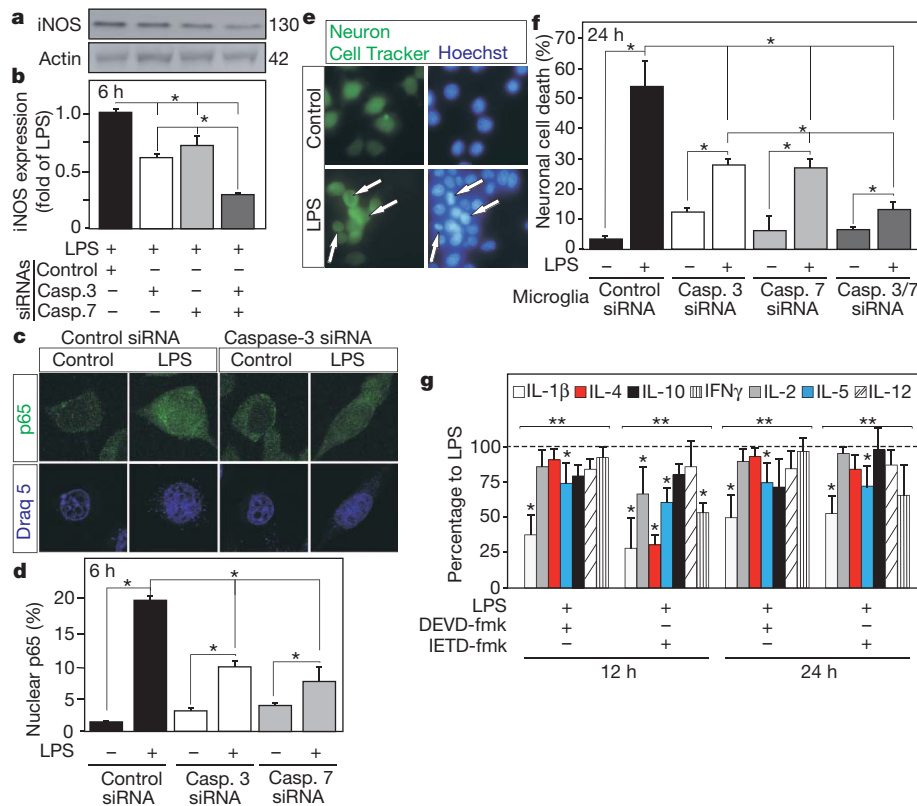


Figure 2 | Knockdown of caspase-3 or caspase-7 decreases microglia activation in response to LPS. siRNAs targeting caspase-3 and/or caspase-7 prevent LPS-induced increase of iNOS expression (a, b) in cultured BV2 cells. In addition, caspase-3 or caspase-7 silencing prevents LPS-induced activation of NF- κ B as seen by reduction of p65 nuclear translocation (c, d). Under co-cultured conditions, LPS treatment induces neuronal cell death and caspase-3/7 knockdown in microglia cells prevents neuronal cell death (e, f). DEVD-fmk or IETD-fmk (caspase-8 inhibitor) treatment modulates cytokine expression at 12 and 24 h in primary microglia cell culture (g). Data are expressed as mean \pm s.e.m. ($n = 3$) and \pm s.d. ($n = 4$) in g. * $P < 0.05$. In g, ** denotes statistically significant difference between treatments ($P < 0.001$) concerning time and treatment.

detect any VEID-ase activity (related to caspase-6) in cells upon LPS treatment. Pretreatment with VEID-fmk, a caspase 6 inhibitor, did not affect LPS-induced DEVD-ase activity, indicating that this protease does not play an essential role in the activation process (Supplementary Fig. 11a, b).

The caspase signalling depends on TLR4

We then examined further the link between TLR4 ligation and activation of caspase-8. Selective knockdown of TLR4 (Supplementary Fig. 12a) was associated with a reduced caspase-8 activation and consequent caspase-3 activation, providing evidence for a direct activation of caspase-8 by TLR4 ligation (Supplementary Fig. 12b, c). Apoptosis after TLR2 activation has been reported to be associated to the formation of a myeloid differentiation factor 88 (MyD88)/Fas-associated

death domain protein (FADD)/Caspase-8 complex²⁶. To assess the potential role of this complex in TLR4 ligation-induced microglia activation, we knocked down MyD88 by siRNAs (Supplementary Fig. 12d). We observed that downregulation of MyD88 did not affect LPS-induced activation of caspase-3 and caspase-8 (Supplementary Fig. 12e, f), suggesting that TLR4 ligation-induced microglia activation did not act through recruitment of the MyD88/FADD/caspase-8, as is the case for TLR2 ligation-induced apoptosis^{26,27}.

Microglia activation can be regulated by autocrine signalling of TNF- α by the TNF receptor 1 (TNFR1) secreted by LPS-stimulated microglia²⁸. To examine whether we were observing a secondary effect, induced by autocrine signalling of TNF- α excreted by LPS-stimulated microglia cells, we investigated the effect of neutralizing TNF- α receptor antibodies on LPS-induced caspase-3 and caspase-8 activities²⁸ (Supplementary Fig. 13a-d). Selectively blocking the TNF- α receptor resulted in decreased LPS-induced iNOS expression at 24 h, thus confirming the TNF- α positive feedback loop exerted on microglia cells (Supplementary Fig. 13e). However, it failed to affect the LPS-induced caspase-3 and caspase-8 activities (Supplementary Fig. 13a-d). Thus, this experiment demonstrates that the caspase-signalling pathway is directly activated, independently of TNF- α receptor stimulation, upon LPS treatment.

Cleaved PKC- δ mediates effect of caspases

We then explored how activated caspase-3/7 interacts with the IKK/NF- κ B pathway during LPS-induced microglia activation. PKC- δ has been reported to regulate NF- κ B activation through the IKK complexes and phosphorylation of the NF- κ B inhibitor I κ B^{29,30}. Interestingly, PKC- δ can be cleaved by caspases to generate a 40 kDa catalytically active fragment³¹. Because the levels of this protein are quite low in this cell line, we decided to overexpress PKC- δ to detect the cleaved form (Supplementary Fig. 14a). We found that LPS treatment of BV2 microglia cells promoted cleavage of PKC- δ into its 40-kDa active fragment (Fig. 4a, b). We found that the caspase inhibitor DEVD-fmk or selective siRNA knockdown of caspase-3 or caspase-7 reduced LPS-induced PKC- δ activation. Therefore, in this context, PKC- δ activation is dependent on DEVD-ase activity (Fig. 4c, d). We then obtained further

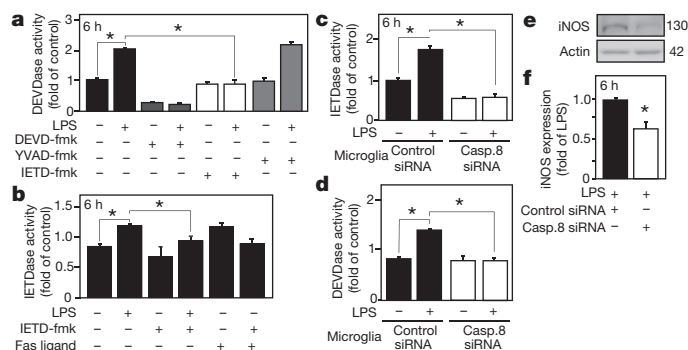


Figure 3 | Caspase-8 activity controls LPS-induced caspase-3/7 activation. Caspase-8 inhibition by IETD-fmk, but not caspase-1 inhibition by YVAD-fmk, prevents LPS-induced DEVDase activity (a). LPS treatment induces caspase-8 (IETDase) activity, which is inhibited by IETD-fmk (b) or specific siRNA targeting caspase-8 (c) in BV2 cells. Treatment with Fas ligand was used as positive control for IETDase activity induction (b). Caspase-8 knockdown using siRNA prevents LPS-induced DEVDase activity (d). siRNA knockdown of caspase-8 prevents LPS-induced expression of iNOS (e, f). Data are expressed as mean \pm s.e.m. ($n = 3$). * $P < 0.05$.

evidence supporting that PKC- δ activation is important in microglia activation. First, we found that the PKC- δ inhibitor rottlerin inhibits LPS-induced iNOS expression in microglia (Fig. 4e, f). Second, if we overexpressed in microglia cells a caspase-uncleavable mutant of PKC- δ , we observed a decrease of LPS-induced iNOS expression (Fig. 4g, h). By contrast, when we overexpressed PKC- δ , we further enhanced the activating effect of LPS on microglia (Supplementary Fig. 14b). Also, we observed a marked decrease of all cytokines in primary microglial cell cultures after being challenged with rottlerin (Fig. 4i). Taken together, our experiments demonstrate that the caspase-3/7-dependent activation of microglia by LPS is mediated through PKC- δ .

Role of the caspase signalling *in vivo*

To examine the physiological relevance of these findings, we performed *in vivo* experiments and injected LPS into the rat substantia nigra. The substantia nigra is known to exhibit a strong inflammatory response upon LPS challenge⁴. At 24 h after injection (Fig. 5a–c), we observed a strong induction of caspase-8 and activation of caspase-3, which was mostly confined to reactive microglia in the mesencephalon on the injected side (Fig. 5b, c). By contrast, microglia were quiescent in the contralateral control midbrain (Fig. 5a). To study if caspase-3/7 activation is important for microglia activation also *in vivo*, we co-injected DEVD-fmk with LPS. Twenty-four hours later, we found that caspase-3/7 inhibition prevented LPS-induced microglia activation (Supplementary Fig. 15a, b). At the molecular level, we observed that the DEVD-fmk treatment mitigated LPS-induced expression of cytokines and pro-inflammatory molecules including iNOS, TNF- α and IL-1 β (Fig. 5d). As a final step, we examined if our earlier *in vitro* findings about the role of PKC- δ in caspase-mediated microglia activation are relevant *in vivo*. We injected LPS intranigally with IETD-fmk, DEVD-fmk or rottlerin. Twenty-four hours later, we dissected the substantia nigra and measured the iNOS protein level. In agreement with our *in vitro* data, we found that inhibition of caspase-8, caspase-3/7 or PKC- δ activities significantly decreased LPS-induced iNOS expression (Fig. 5e, f).

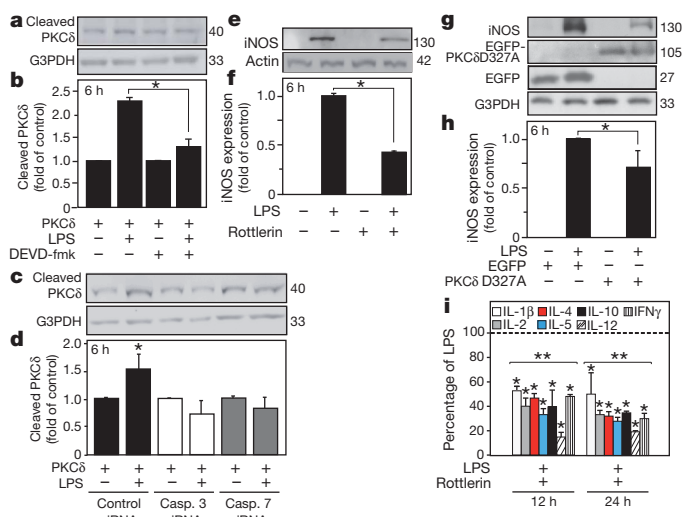


Figure 4 | Caspase-3/7 regulates microglia activation through the PKC- δ pathway. Caspase-3/7 inhibition by DEVD-fmk (**a**, **b**) or by siRNA knockdown (**c**, **d**) prevents LPS-induced PKC- δ cleavage/activation in BV2 microglia cells. PKC- δ inhibition by rottlerin mostly abolishes LPS-induced activation of iNOS (**e**, **f**). Overexpression of PKC- δ D327A, a caspase uncleavable PKC- δ mutant, has an inhibitory effect on LPS-induced iNOS expression (**g**, **h**). Rottlerin treatment reduces LPS-induced cytokine expression at 12 and 24 h in primary microglia cell culture (**i**). Data are expressed as mean \pm s.e.m. ($n = 3$) and \pm s.d. ($n = 4$) in **i**. * $P < 0.05$. In **i**, ** denotes statistically significant difference between treatments $P < 0.001$, concerning time and treatment.

To substantiate further our observations in the rat LPS model, we monitored the activation of microglia in the MPTP-lesion mouse model of PD. Consistent with earlier work, mice injected systemically with MPTP showed a strong microglial activation and a robust neurotoxicity in ventral mesencephalic dopamine neurons³². Intranigral vehicle injections in sham animals (saline containing 1% DMSO) greatly increased the density of reactive microglia, as evidenced by Iba1 immunohistochemistry (Supplementary Fig. 15c, d). Consequently, we quantified the numbers of reactive and resting microglia in Iba1-immunostained sections covering injection and non-injection sites for each experimental condition (Fig. 5g, h and Supplementary Fig. 15c, d). As expected, the density of reactive microglia at the non-injection site was highest in the MPTP group and dramatically lower in the sham (no MPTP) group. Interestingly, caspase 8 inhibition significantly prevented the MPTP-induced microglia activation and the MPTP-induced reduction in the density of resting microglia (Fig. 5g, h). Similar results of IETD-fmk were observed at the injection site (Supplementary Fig. 14c, d). As expected, MPTP injections severely reduced the integrity of nigro-striatal dopaminergic terminals (Supplementary Fig. 15e–g). Intranigral IETD-fmk injection induced a modest, but significant, protection against MPTP-induced toxicity as demonstrated by densitometric analysis of surviving striatal dopaminergic terminals and stereological cell counts in the substantia nigra (Supplementary Fig. 15e–g).

Caspase activation in human brains

PD and AD are known to be associated with neuroinflammation and the presence of activated microglia^{2,33}. We investigated whether caspase-3 and caspase-8 are activated in microglia of individuals with PD and AD. We analysed expression of cleaved caspase-3, cleaved caspase-8 and the microglia marker CD68 in post-mortem brain samples from individuals with PD and AD (clinically as well as neuropathologically diagnosed) and age- and gender-matched healthy control brains. We detected significant cytoplasmic expression of both active caspase-3 and active caspase-8 in the PD ventral mesencephalon and in the AD frontal cortex, compared with controls (Fig. 6a–d and Supplementary Figs 16–19). The activated caspases and the microglial marker CD68 were largely co-localized, indicating that caspase-8 and caspase-3 are activated mainly in microglia in PD and AD (Fig. 6a–d and Supplementary Fig. 20a, b).

Discussion

In this study, we uncover a completely novel role for caspase-8 and caspase-3/7 in the control of microglia and brain inflammation. We show that stimulation of microglia with the pro-inflammatory stimuli LPS, LTA, PamC3sk4 and IFN- γ triggers caspase-3/7 activation, without causing cell death. Caspases are proteases essential for apoptosis and inflammation. Caspase-1 is already known as the prototypical inflammatory caspase, required for the maturation of pro-inflammatory cytokines. By comparison, the initiator caspase-8 and effector pro-apoptotic caspase-3 and caspase-7 are considered crucial in the intracellular death machinery. Unexpectedly, we showed that caspase-3/7-dependent DEVD-ase activity controls LPS-induced microglia activation. We demonstrated that inhibition of the caspase-3/7 pathway effectively block microglia activation. For example, we observed fewer microglia with an activated phenotype in the presence of the caspase inhibitor DEVD-fmk and inhibition of the downstream IKK/NF- κ B pathways. We also reveal that caspase-3/7 activates the IKK/NF- κ B pathways through processing and activation of PKC- δ . Furthermore, we found that microglia exposed to the pro-inflammatory agent LPS failed to be toxic to neighbouring neurons when we inhibited caspase-3/7 chemically or by siRNA gene silencing. Importantly, we provide compelling evidence that active caspases 3 and 8 are expressed within reactive microglia in the ventral mesencephalon and frontal cortex of individuals with PD and AD, respectively.

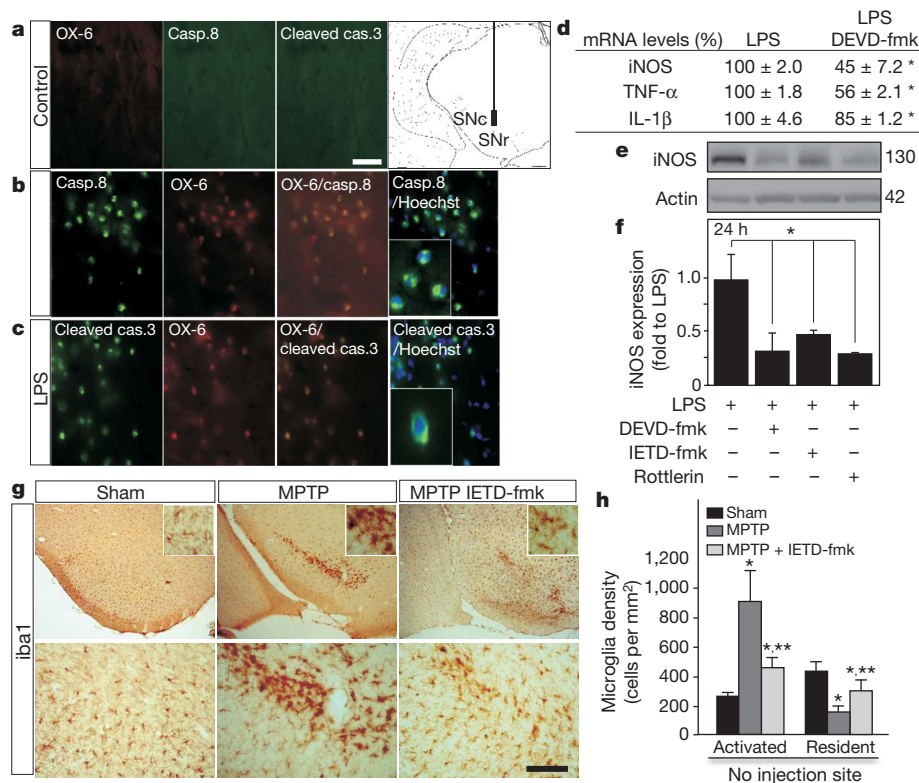


Figure 5 | *In vivo* inhibition of the caspase-dependent pathways prevents microglia activation. *In vivo* injection of LPS in rat substantia nigra (insert in **a**) induced activation of caspase-8 and caspase-3 in the ventral mesencephalon, which is mostly localized in OX-6-labelled reactive microglia (**b**, **c**) compared with control (**a**). There is a null coincidence between Hoechst-labelled nuclei and caspase8/cleaved caspase 3 immunoreactivity in OX-6-labelled microglia (**b**, **c**). *In vivo* DEVD-fmk significantly reduced intranigral LPS-induced microglia activation as evaluated by mRNA levels of proinflammatory molecules (**d**). *In vivo* inhibition of caspase-8, caspase-3/7 or PKC-δ, using IETD-fmk, DEVD-fmk and Rottlerin respectively, predominantly abolished the intranigral induction of iNOS upon LPS treatment (**e**, **f**). *In vivo* acute intraperitoneal injections of MPTP in mice highly increased the density of reactive microglia in substantia nigra (**g**, **h**). Intranigral IETD-fmk injections robustly prevented the MPTP-induced microglia activation and decrease of the resident microglia density at the non-injection site (**g**, **h**). Data are expressed as mean ± s.e.m. ($n = 3$). SNr, substantia nigra pars reticulata; SNC, Substantia nigra pars compacta. Scale bars: **a–c**, 25 μm; **d**, 500 μm for low magnification photographs; **j**, 400 and 100 μm for low- and high-magnification photographs. * $P < 0.05$; ** denotes statistically significant difference between MPTP in **h**, and MPTP+IETD-fmk treatments.

To summarize, we present new non-apoptotic functions for caspases-8, -3 and -7 and show that they can have a pivotal role in inflammation of the central nervous system (Supplementary Fig. 1). Brain inflammation is a typical feature of neurodegenerative diseases^{1,2,34} and is a prominent sequel of many acute forms of brain injury (for example, trauma, encephalitis and stroke)^{35,36}. Under certain circumstances, neuroinflammation is known to promote neuronal death³⁷. Accordingly, previous studies have shown that

anti-inflammatory treatment can reduce neurodegeneration. Our discovery that the caspases-8, -3 and -7 cascade can promote neuroinflammation through IKK/NF-κB and PKC-δ, with the development of nanocarriers that allow the inhibitor to cross the blood–brain barrier, opens up new molecular targets for anti-inflammatory drugs³⁸. Considering that these caspases also regulate apoptotic neuronal death, our results should revitalize interest in caspase inhibitors as potential therapeutic agents in disorders of the central nervous system^{39–41}.

METHODS SUMMARY

Microglial BV2 and dopaminergic MN9D murine cells were cultured as described^{42,43}. Experiments were performed in reduced 5% FCS media. Transfection of BV2 cells used Lipofectamine 2000 (Invitrogen). Non-targeting control, caspase-3, caspase-7 and caspase-8 siRNA were obtained from Dharmacon. Single and co-cultures were exposed to 1 μg ml⁻¹ LPS. Primary microglial cells were prepared from postnatal day 1–3 mouse brain using a previously described protocol⁴⁴. Cytokine content was assayed using the Mouse TH1/TH2 9-PlexTissue Culture Kit (Meso-scale Discovery). DEVD-fmk (20 μM), IETD-fmk (20 μM), YVAD-fmk (20 μM), rottlerin (2 μM) or TNF R1 mouse Ab (20 ng ml⁻¹) were added to the media 1 h before LPS treatment. Under approved protocols, male albino Wistar rats (230–250 g) were intranigally injected with 2 μg LPS alone or in combination with 0.75 nmol DEVD-fmk or 0.75 nmol IETD-fmk. Rottlerin (20 mg kg⁻¹) was administered intraperitoneally. Twenty-four hours after surgery, rats were killed and brains processed for analysis. Male C57BL mice were intranigally injected with vehicle or 0.75 nmol IETD-fmk. Twelve hours later, animals were treated with four injections of MPTP (16 mg kg⁻¹) at 2-h intervals. Four days after the last injection, animals were killed and brains processed for analysis. Paraffin-embedded archival tissue blocks from autopsy on three individuals with PD and three individuals with AD were obtained from the Department of Pathology, Lund University Hospital, Sweden. Age-matched controls (cardiac arrest victims, no brain disease) were analysed at the same time. Histological, immunological and fluorescence-activated cell sorting (FACS) analyses using antibodies listed in Supplementary Table 1 were performed using standard procedures^{42,45–47}. The subcellular localization of cleaved caspase-3, cleaved caspase-8 or NF-κB p65 protein was determined by confocal microscopy. Quantitative PCR with primers listed in Supplementary Table 2 were performed using SensiMixPlus SYBR (Quantace). MTT, caspase activity assay, apoptosis quantification and FACS analysis have been previously described⁴⁵. Statistical evaluations were performed by one- or two-way analysis of variance with Bonferroni post hoc tests and Kolmogorov–Smirnov test.

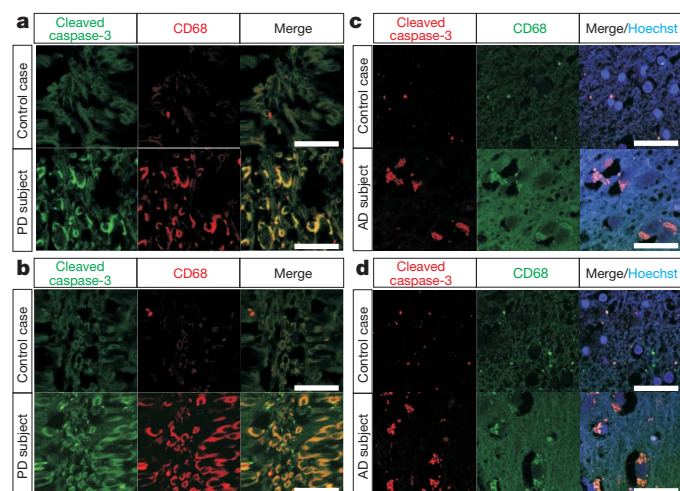


Figure 6 | Activation of caspase-3 and caspase-8 in microglia in brain from individuals with PD and AD. Double immunolabelling and confocal imaging analysis show significant activation of caspase-8 and caspase-3, mostly localized in CD68-labelled microglia, in the ventral mesencephalon from patients with PD (**a**, **b**) and in the frontal cortex of patients with AD (**c**, **d**) compared with age- and gender-matched healthy controls. Confocal images of immunostaining with cleaved caspase-3 (**a**, **c**) or cleaved caspase-8 (**b**, **d**), together with CD68 antibodies, are depicted. Merge from a control and an individual with AD illustrates the absence of cleaved caspase 3/caspase staining in Hoechst-labelled nuclei in reactive microglia (**c**, **d**). Scale bar, 10 μm.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 October; accepted 23 December 2010.

Published online 9 March 2011.

1. Hanisch, U. K. & Kettenmann, H. Microglia: active sensor and versatile effector cells in the normal and pathologic brain. *Nature Neurosci.* **10**, 1387–1394 (2007).
2. Block, M. L., Zecca, L. & Hong, J. S. Microglia-mediated neurotoxicity: uncovering the molecular mechanisms. *Nature Rev. Neurosci.* **8**, 57–69 (2007).
3. Chao, C. C., Hu, S., Molitor, T. W., Shaskan, E. G. & Peterson, P. K. Activated microglia mediate neuronal cell injury via a nitric oxide mechanism. *J. Immunol.* **149**, 2736–2741 (1992).
4. Castano, A., Herrera, A. J., Cano, J. & Machado, A. Lipopolysaccharide intranigral injection induces inflammatory reaction and damage in nigrostriatal dopaminergic system. *J. Neurochem.* **70**, 1584–1592 (1998).
5. Saijo, K. *et al.* A Nurr1/CoREST pathway in microglia and astrocytes protects dopaminergic neurons from inflammation-induced death. *Cell* **137**, 47–59 (2009).
6. Zhao, J. *et al.* IRF-8/interferon (IFN) consensus sequence-binding protein is involved in Toll-like receptor (TLR) signaling and contributes to the cross-talk between TLR and IFN-gamma signaling pathways. *J. Biol. Chem.* **281**, 10073–10080 (2006).
7. Car, B. D. *et al.* Interferon gamma receptor deficient mice are resistant to endotoxic shock. *J. Exp. Med.* **179**, 1437–1444 (1994).
8. Jin, J. J., Kim, H. D., Maxwell, J. A., Li, L. & Fukuchi, K. Toll-like receptor 4-dependent upregulation of cytokines in a transgenic mouse model of Alzheimer's disease. *J. Neuroinflammation* **5**, 23 (2008).
9. Balistreri, C. R. *et al.* Association between the polymorphisms of TLR4 and CD14 genes and Alzheimer's disease. *Curr. Pharm. Des.* **14**, 2672–2677 (2008).
10. Walter, S. *et al.* Role of the toll-like receptor 4 in neuroinflammation in Alzheimer's disease. *Cell. Physiol. Biochem.* **20**, 947–956 (2007).
11. Nicholson, D. W. *et al.* Identification and inhibition of the ICE/CED-3 protease necessary for mammalian apoptosis. *Nature* **376**, 37–43 (1995).
12. Cohen, G. M. Caspases: the executioners of apoptosis. *Biochem. J.* **326**, 1–16 (1997).
13. Keller, M., Ruegg, A., Werner, S. & Beer, H. D. Active caspase-1 is a regulator of unconventional protein secretion. *Cell* **132**, 818–831 (2008).
14. Schulz, J. B. *et al.* Extended therapeutic window for caspase inhibition and synergy with MK-801 in the treatment of cerebral histotoxic hypoxia. *Cell Death Differ.* **5**, 847–857 (1998).
15. Braun, J. S. *et al.* Neuroprotection by a caspase inhibitor in acute bacterial meningitis. *Nature Med.* **5**, 298–302 (1999).
16. Cutillas, B., Espejo, M., Gil, J., Ferrer, I. & Ambrosio, S. Caspase inhibition protects nigral neurons against 6-OHDA-induced retrograde degeneration. *Neuroreport* **10**, 2605–2608 (1999).
17. Depino, A. M. *et al.* Microglial activation with atypical proinflammatory cytokine expression in a rat model of Parkinson's disease. *Eur. J. Neurosci.* **18**, 2731–2742 (2003).
18. Kawai, T. & Akira, S. Signaling to NF- κ B by Toll-like receptors. *Trends Mol. Med.* **13**, 460–469 (2007).
19. Gibbons, H. M. & Dragunow, M. Microglia induce neural cell death via a proximity-dependent mechanism involving nitric oxide. *Brain Res.* **1084**, 1–15 (2006).
20. Schumann, R. R. *et al.* Lipopolysaccharide activates caspase-1 (interleukin-1-converting enzyme) in cultured monocytic and endothelial cells. *Blood* **91**, 577–584 (1998).
21. Li, P. *et al.* Mice deficient in IL-1 β -converting enzyme are defective in production of mature IL-1 β and resistant to endotoxic shock. *Cell* **80**, 401–411 (1995).
22. Friedlander, R. M. *et al.* Expression of a dominant negative mutant of interleukin-1 β converting enzyme in transgenic mice prevents neuronal cell death induced by trophic factor withdrawal and ischemic brain injury. *J. Exp. Med.* **185**, 933–940 (1997).
23. Fernandes-Alnemri, T. *et al.* *In vitro* activation of CPP32 and Mch3 by Mch4, a novel human apoptotic cysteine protease containing two FADD-like domains. *Proc. Natl Acad. Sci. USA* **93**, 7464–7469 (1996).
24. Nunez, G., Benedict, M. A., Hu, Y. & Inohara, N. Caspases: the proteases of the apoptotic pathway. *Oncogene* **17**, 3237–3245 (1998).
25. Slee, E. A., Adrain, C. & Martin, S. J. Serial killers: ordering caspase activation events in apoptosis. *Cell Death Differ.* **6**, 1067–1074 (1999).
26. Aliprantis, A. O., Yang, R. B., Weiss, D. S., Godowski, P. & Zychlinsky, A. The apoptotic signaling pathway activated by Toll-like receptor-2. *EMBO J.* **19**, 3325–3336 (2000).
27. Jung, D. Y. *et al.* TLR4, but not TLR2, signals autoregulatory apoptosis of cultured microglia: a critical role of IFN-beta as a decision maker. *J. Immunol.* **174**, 6467–6476 (2005).
28. Kuno, R. *et al.* Autocrine activation of microglia by tumor necrosis factor-alpha. *J. Neuroimmunol.* **162**, 89–96 (2005).
29. Storz, P., Doppler, H. & Toker, A. Protein kinase Cdelta selectively regulates protein kinase D-dependent activation of NF- κ B in oxidative stress signaling. *Mol. Cell. Biol.* **24**, 2614–2626 (2004).
30. Vancurova, I., Miskolci, V. & Davidson, D. NF- κ B activation in tumor necrosis factor α -stimulated neutrophils is mediated by protein kinase C δ . Correlation to nuclear I κ B α . *J. Biol. Chem.* **276**, 19746–19752 (2001).
31. Reyland, M. E., Anderson, S. M., Matassa, A. A., Barzen, K. A. & Quissell, D. O. Protein kinase C δ is essential for etoposide-induced apoptosis in salivary gland acinar cells. *J. Biol. Chem.* **274**, 19115–19123 (1999).
32. Czlonkowska, A., Kohutnicka, M., Kurkowska-Jastrzebska, I. & Czlonkowski, A. Microglial reaction in MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) induced Parkinson's disease mice model. *Neurodegeneration* **5**, 137–143 (1996).
33. Aarli, J. A. Role of cytokines in neurological disorders. *Curr. Med. Chem.* **10**, 1931–1937 (2003).
34. Gonzalez-Scarano, F. & Baltuch, G. Microglia as mediators of inflammatory and degenerative diseases. *Annu. Rev. Neurosci.* **22**, 219–240 (1999).
35. Jordan, J., Segura, T., Brea, D., Galindo, M. F. & Castillo, J. Inflammation as therapeutic objective in stroke. *Curr. Pharm. Des.* **14**, 3549–3564 (2008).
36. Lenzlinger, P. M., Morganti-Kossmann, M. C., Laurer, H. L. & McIntosh, T. K. The duality of the inflammatory response to traumatic brain injury. *Mol. Neurobiol.* **24**, 169–181 (2001).
37. Allan, S. M. & Rothwell, N. J. Inflammation in central nervous system injury. *Phil. Trans. R. Soc. Lond. B* **358**, 1669–1677 (2003).
38. Karatas, H. *et al.* A nanomedicine transports a peptide caspase-3 inhibitor across the blood-brain barrier and provides neuroprotection. *J. Neurosci.* **29**, 13761–13769 (2009).
39. Bilsland, J. & Harper, S. Caspases and neuroprotection. *Curr. Opin. Investig. Drugs* **3**, 1745–1752 (2002).
40. Friedlander, R. M. Apoptosis and caspases in neurodegenerative diseases. *N. Engl. J. Med.* **348**, 1365–1375 (2003).
41. Le, D. A. *et al.* Caspase activation and neuroprotection in caspase-3-deficient mice after *in vivo* cerebral ischemia and *in vitro* oxygen glucose deprivation. *Proc. Natl Acad. Sci. USA* **99**, 15188–15193 (2002).
42. Joseph, B. *et al.* p57(Kip2) cooperates with Nurr1 in developing dopamine cells. *Proc. Natl Acad. Sci. USA* **100**, 15619–15624 (2003).
43. Bocchini, V. *et al.* An immortalized cell line expresses properties of activated microglial cells. *J. Neurosci. Res.* **31**, 616–621 (1992).
44. Giulian, D. & Baker, T. J. Characterization of ameboid microglia isolated from developing mammalian brain. *J. Neurosci.* **6**, 2163–2178 (1986).
45. Li, J. Y. *et al.* Lewy bodies in grafted neurons in subjects with Parkinson's disease suggest host-to-graft disease propagation. *Nature Med.* **14**, 501–503 (2008).
46. Joseph, B. *et al.* Mitochondrial dysfunction is an essential step for killing of non-small cell lung carcinomas resistant to conventional treatment. *Oncogene* **21**, 65–77 (2002).
47. Rite, I., Machado, A., Cano, J. & Venero, J. L. Blood-brain barrier disruption induces *in vivo* degeneration of nigral dopaminergic neurons. *J. Neurochem.* **101**, 1567–1582 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Gorman, O. Hermanson, M. Malewicz, S. Orrenius, T. Panaretakis and B. Zhivotovsky for discussion, and L. Hjortsberg, M. Reyland and S. Ceccatelli for providing us with reagents. M. Carballo, J.L. Ribas, A. Fernández and B. Haraldsson provided qualified technical support. This work has been supported by grants from the Spanish Ministerio de Ciencia y Tecnología (SAF2006-04119 and 2009-13778), the Swedish Research Council, the Parkinson Foundation of Sweden, the Swedish Alzheimer Foundation and the Swedish Cancer Society. M.A.B., T.D. and P.B. are members of Neurofortis and Bagadilico, both of which are research environments sponsored by the Swedish Research Council.

Author Contributions M.A.B. performed all the experiments except as otherwise noted. qPCR was performed by A.G.-Q. and E.K. J.L.V. and J.C. collaborated in doing surgery and further dissecting the animal brains. M.A.B. and T.D. performed primary cell culture experiments and cytokine analysis. E.K. collaborated in performing the caspase activity assay. B.J. and E.K. collaborated in performing FACS. B.J. collaborated also in the confocal imaging analysis. E.E. did the neuropathology of the individuals with PD and AD and the controls. A.P. prepared tissue and participated in the morphological assessment of human brain specimens. N.H. and P.B. were involved in study design. M.A.B., J.L.V. and B.J. designed the study, analysed and interpreted the data. All authors discussed the results and commented on or edited the manuscript. The first draft of the paper was written by B.J. J.L.V. and B.J. share senior authorship of the paper. T.D. and E.K. share second authorship.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.J. (bertrand.joseph@ki.se) or J.L.V. (jlvenero@us.es).

METHODS

Reagents. LPS (from *Escherichia coli*, serotype 026:B6; Sigma), staurosporine (STS; Sigma), MPTP (Sigma), agonist anti-Fas monoclonal antibody clone CH11 (MBL), agonist anti-TNF- α (R&D Systems), the caspase-3/7 inhibitor benzyloxycarbonyl-Asp(OMe)-Glu(OMe)-Val-Asp(OMe)-fluoromethylketone (DEVD-fmk), the caspase-8 inhibitor benzyloxycarbonyl-Ile-Glu(OMe)-Thr-Asp(OMe)-fluoromethylketone (IETD-fmk), the caspase-6 inhibitor benzyloxycarbonyl-Val-Glu(OMe)-Ile-Asp(OMe)-fluoromethylketone (VEID-fmk), the caspase-1 inhibitor benzyloxycarbonyl-Tyr-Val-Ala-Asp(OMe)-fluoromethylketone (YVAD-fmk; MP Biomedicals) and the PKC- δ inhibitor Rottlerin (Calbiochem) were used in this study. Plasmids encoding PKC- δ and EGFP-PKC- δ D327A were gifts of L. Hjörstberg and M. Reyland respectively. ON-TARGET plus SMARTpools siRNAs were purchased from Dharmacon (Supplementary Table 3).

Animals and surgery. Animals used in this study were obtained from the Center of Production and Animal Experimentation (Estartinas, Spain) and NMRI (Charles River, Germany). Experiments were performed in accordance with the Guidelines of the European Union Council (86/609/EU), following Spanish and Swedish regulations for the use of laboratory animals and approved by the Scientific Committee of the University of Seville and Lund University. Intraneural injections were made 5.8 mm anterior, 2.0 mm lateral and 8.0 mm ventral to the bregma in rat, and 3.1 mm anterior, \pm 1.2 mm lateral and 5.1 mm ventral from bregma in mouse.

Co-culture and neuronal cell death assay. Microglial BV2 and dopaminergic MN9D murine cells were cultured as described^{42,43}. MN9D dopaminergic neuronal cells were stained with CellTracker Green CMFDA (Invitrogen) before BV2 microglia cells transfected with caspase-3 siRNA, caspase-7 siRNA or non-targeting siRNA were plated on them. After 24 h, cells were treated with LPS and incubated for an additional 24 h. They were then fixed with 4% paraformaldehyde and stained with 0.1 mg ml⁻¹ Hoechst for quantification of neuronal cell death.

Primary cultures. Primary mouse microglial cells were plated for at least 48 h before the experiments (ethical permit M302-09). Cells were pretreated with inhibitors (20 μ M DEVD-fmk, 20 μ M IETD-fmk and 2 μ M Rottlerin 1 h before with 100 ng ml⁻¹ LPS treatment)⁴⁴.

Human brain. Human brain tissues from patients with PD of 5, 9 and 15 years' duration, respectively, and age-matched controls were used in this study⁴⁵. The region investigated was the anterior mesencephalon covering the substantia nigra. Frontal cortex from patients with AD of 4, 10 and 14 years' duration and controls were also used (Regional Ethical Review Board, Lund (Sweden) 2009-646/2010-25). The patients with dissimilar disease duration exhibited different degrees of severity of brain disease, reflecting different stages of the degenerative process. All sections were stained with haematoxylin-erythrosin and with antibodies against caspase 3, 8 and microglia, CD68 (see below). They were microscopically reviewed for verification of pathology. Before the investigation, the entire collection of brain sections, 15–20 per individual including the mesencephalic section, were subjected to a neuropathological whole-brain analysis for clinical diagnostic purposes, according to routine procedures at the Department of Pathology, Division of Neuropathology, Lund University Hospital. The project procedures involving human brain tissue were approved by the Regional Ethical Review Board in Lund, Sweden.

Immunofluorescence and laser scanning confocal microscopy. Paraformaldehyde-fixed cells were blocked in PBS/3% goat serum/0.3% triton X-100 and incubated

with the indicated primary (4 °C, overnight) and AlexaFluor 488 or 594 conjugated anti-IgG used as secondary antibodies (room temperature, ~22 °C, 1 h; Molecular Probes). Nuclei were counterstained with DRAQ5TM (1 μ M, Alexis), Hoechst (1 μ g ml⁻¹, Molecular Probes) or DAPI (1 μ g ml⁻¹, Molecular Probes). Alexa-555-conjugated cholera toxin B (CTB; 10 μ g ml⁻¹; Molecular Probes) was used to stain lipid rafts on the plasma membrane. Protein subcellular localization was analysed under a Zeiss 510 META or Leica TCS-SP2 confocal laser scanning microscope. The nuclear translocation of the NF- κ B p65 subunit was quantified using multicolour three-dimensional plug-in from Leica Confocal software and measured as the percentage of co-localization of DRAQ5TM with NF- κ B p65 inside the nucleus.

FACS analysis. Quantification of cells with cleaved caspase-3 and cleaved caspase-8 was performed with a FACSCalibur flow cytometer (Becton Dickinson) using standard procedures. Analysis of data was performed using Cell Quest software⁴⁶.

Immunohistochemistry and immunohistochemistry. Rats and mice were perfused through the heart under deep anaesthesia with 4% paraformaldehyde/PBS, pH 7.4. Brains were removed, cryoprotected in sucrose and frozen in isopentane at -15 °C; serial coronal sections (25- and 30- μ m sections, for rat and mice, respectively) covering the striatum and the substantia nigra were cut with a cryostat and mounted on gelatin-coated slides⁴⁷. Sections were incubated with the indicated primary antibodies. After three washes, sections were incubated with biotinylated horse anti-mouse or goat anti-rabbit IgG (Vector) followed by an incubation with ExtrAvidin-Peroxidase solution (Sigma) and for immunofluorescence by a fluorescein isothiocyanate (FITC)-conjugated anti-rabbit and Texas Red anti-mouse antibody (Vector). The peroxidase was visualized with a standard diaminobenzidine/hydrogen reaction for 5 min. For paraffin-embedded human tissue material, sections (5 μ m) were mounted on capillary glass slides (DAKO). Sections were microwaved pre-treated in 10 mM citrate buffer pH 6.0 for 15 min at 800 W for antigen retrieval. An automated immunostainer (TechMateTM 500 Plus, DAKO) was used for the staining procedure using DAKO ChemMate Kit Peroxidase/3-3'-diaminobenzidine. The indicated primary antibodies were used.

Quantification of microglial population in animal models. Reactive and resident microglial cells were counted in LPS-injected rats and MPTP-exposed mice detected by Iba1 immunohistochemistry based on morphological features. For each animal, eight sections covering the entire antero-posterior ventral mesencephalon were analysed. For each section, four photographs were taken at \times 20 magnification (two for each substantia nigra) and microglial cells were counted with the analySIS software.

Quantification of the striato-nigral dopaminergic system. Tyrosine hydroxylase immunohistochemistry in the striatum of mice intoxicated with MPTP was quantified using a computer-assisted software (analySIS). For the quantification, five striatal sections from each condition, processed under identical experimental conditions, were scanned at high resolution. The striatal region was delineated and its optical density measured based upon a calibrated grey scale. Quantification of tyrosine-hydroxylase-positive cells in the substantia nigra was performed according to a modified stereological approach using the Olympus CAST-Grid system. The area of the substantia nigra region was estimated using the principle of Cavalieri. All data were collected while blind to experimental treatment and expressed as number of neurons per substantia nigra.

Crystal structure of metarhodopsin II

Hui-Woog Choe^{1,2*}, Yong Ju Kim^{1*}, Jung Hee Park^{1*}, Takefumi Morizumi^{1†}, Emil F. Pai³, Norbert Krauß⁴, Klaus Peter Hofmann^{1,5}, Patrick Scheerer¹ & Oliver P. Ernst^{1†}

G-protein-coupled receptors (GPCRs) are seven transmembrane helix (TM) proteins that transduce signals into living cells by binding extracellular ligands and coupling to intracellular heterotrimeric G proteins (Gαβγ)¹. The photoreceptor rhodopsin couples to transducin and bears its ligand 11-*cis*-retinal covalently bound via a protonated Schiff base to the opsin apoprotein². Absorption of a photon causes retinal *cis/trans* isomerization and generates the agonist all-*trans*-retinal *in situ*. After early photoproducts, the active G-protein-binding intermediate metarhodopsin II (Meta II) is formed, in which the retinal Schiff base is still intact but deprotonated. Dissociation of the proton from the Schiff base breaks a major constraint in the protein and enables further activating steps, including an outward tilt of TM6 and formation of a large cytoplasmic crevice for uptake of the interacting C terminus of the Gα subunit^{3–5}. Owing to Schiff base hydrolysis, Meta II is short-lived and notoriously difficult to crystallize. We therefore soaked opsin crystals with all-*trans*-retinal to form Meta II, presuming that the crystal's high concentration of opsin in an active conformation (Ops*)^{6,7} may facilitate all-*trans*-retinal uptake and Schiff base formation. Here we present the 3.0 Å and 2.85 Å crystal structures, respectively, of Meta II alone or in complex with an 11-amino-acid C-terminal fragment derived from Gα (GαCT2). GαCT2 binds in a large crevice at the cytoplasmic side, akin to the binding of a similar Gα-derived peptide to Ops* (ref. 7). In the Meta II structures, the electron density from the retinal ligand seamlessly continues into the Lys 296 side chain, reflecting proper formation of the Schiff base linkage. The retinal is in a relaxed conformation and almost undistorted compared with pure crystalline all-*trans*-retinal. By comparison with early photoproducts we propose how retinal translocation and rotation induce the gross conformational changes characteristic for Meta II. The structures can now serve as models for the large GPCR family.

To obtain crystals of Meta II alone or in complex with the GαCT2 peptide, we followed previous approaches using low pH to obtain crystals of opsin in the active Ops* conformation^{6,7} for subsequent soaking with all-*trans*-retinal. GαCT2 was derived from the transducin Gα C terminus and contained two substitutions which improved crystal diffraction quality compared with crystals containing the native peptide fragment (see Methods). Reversible formation of Meta II is indicated by (1) the 380-nm absorption obtained before and after dissolving the all-*trans*-retinal-soaked opsin or opsin-GαCT2 crystals, respectively, (2) formation of an opsin-retinal Schiff base bond as determined by acid denaturation of the reconstituted pigments (Supplementary Fig. 1), (3) the X-ray crystallographic evidence for the Schiff base bond between all-*trans*-retinal and Lys 296^{7,43} of opsin (see below), and (4) the high activity of the reconstituted pigment towards the G protein (Supplementary Fig. 2). In addition, soaking of the opsin crystals with all-*trans*-retinal for an optimized soaking time had only little effect on the diffraction quality of the opsin crystals,

indicating that the Ops* conformation present in the crystal is compatible with the all-*trans*-retinal agonist. In contrast, soaking of opsin crystals with the inverse agonist 11-*cis*-retinal led to a reddish colour indicative of rhodopsin formation. Soaking was accompanied by severe loss of crystal diffraction quality, indicating disintegration of crystal packing, consistent with the expected larger conformational changes induced in the protein by 11-*cis*-retinal.

The structures of Meta II and Meta II-GαCT2 were solved at 3.0 and 2.85 Å resolution, respectively, by molecular replacement techniques using opsin in the Ops* conformation (PDB accession 3DQB) as search model. For crystallization, data collection, structure determination and refinement statistics, see Methods and Supplementary Table 1. The structures comprise amino acids 1–326 of the opsin protein moiety, lacking 22 C-terminal amino acids, which are not resolved in the corresponding electron density maps, presumably because of high mobility^{6,7}. Each model also includes one molecule all-*trans*-retinal, oligosaccharyl chains at Asn 2^{NT} (for superscripts see Methods Summary) and Asn 15^{NT}, several water and two *n*-octyl-β-D-glucopyranoside molecules, and in addition in the case of Meta II-GαCT2, the 11-amino-acid GαCT2 peptide. The models further comprise one molecule palmitic acid covalently bound to Cys 323^{H8}; a second palmitic acid residue at Cys 322^{H8} is lacking in each model because of weak electron density. The overall structure of Meta II and Meta II-GαCT2, comprising seven TMs connected by cytoplasmic and extracellular loops (CL1–CL3 and EL1–EL3) which are followed by the cytoplasmic helix H8, is similar to the previously solved Ops* and Ops*-GαCT structures^{6,7}, but quite different from inactive rhodopsin^{8–10}. The Cα root mean squared deviation (r.m.s.d.) values in Å are 0.51 (Ops*/Meta II), 0.44 (Ops*-GαCT/Meta II-GαCT2), 0.35 (Meta II/Meta II-GαCT2) and 2.85 (Rhodopsin/Meta II) (Fig. 1 and Supplementary Fig. 3)^{6,7,9}. Both Meta II and Meta II-GαCT2 display the open cytoplasmic receptor surface: due to a rotational tilt of TM6 and motion of TM5 a crevice in the cytoplasmic surface is created into which the Gα C terminus or its mimic, the GαCT2 peptide, can bind^{4,5,7}.

The Meta II structures now provide insight into retinal-protein interactions in the ligand binding site. The initial electron density map calculated with omitted retinal ligand could readily be interpreted with the relaxed all-*trans*-retinal conformation represented by the model obtained from the Cambridge Structural Database (reference code TRETAL02; Supplementary Figs 4 and 5). There is strong electron density connecting the retinal into the side chain of Lys 296^{7,43} consistent with the presence of a Schiff base bond between carbon atom C15 of retinal and the ε-amino group of Lys 296^{7,43} (Fig. 2 and Supplementary Fig. 4). The refinement of the Meta II structures only slightly changed the initial conformation of all-*trans*-retinal, indicating that retinal in a relaxed conformation is bound in the ligand binding site of Ops*, in agreement with spectroscopic studies on Meta II (ref. 11). Both structures, Meta II and Meta II-GαCT2, show an almost identical retinal conformation (Supplementary Fig. 5). The

¹Institut für Medizinische Physik und Biophysik - CC2, Charité - Universitätsmedizin Berlin, Charitéplatz 1, D-10117 Berlin, Germany. ²Department of Chemistry, College of Natural Science, Chonbuk National University, 561-756 Chonju, South Korea. ³Departments of Biochemistry, Molecular Genetics and Medical Biophysics, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada. ⁴Queen Mary University of London, School of Biological and Chemical Sciences, London E1 4NS, UK. ⁵Zentrum für Biophysik und Bioinformatik, Humboldt-Universität zu Berlin, Invalidenstrasse 42, D-10115 Berlin, Germany. [†]Present addresses: Department of Biochemistry, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada (T.M.); Departments of Biochemistry and Molecular Genetics, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada (O.P.E.).

*These authors contributed equally to this work.

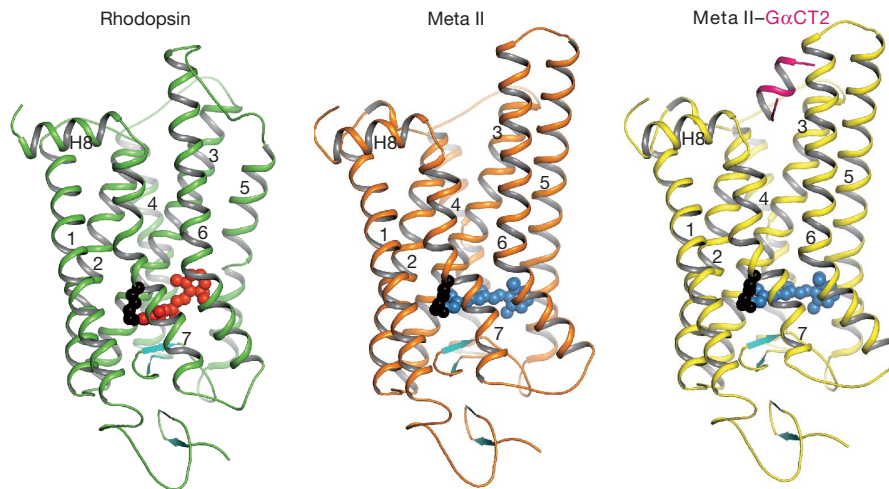


Figure 1 | Structures of inactive rhodopsin, active Meta II and Meta II in complex with a Ga fragment. Side view of cartoon model of rhodopsin (PDB accession 1U19), Meta II and Meta II-GαCT2, that is, Meta II in complex with an 11-amino-acid C-terminal peptide derived from the transducin Gα subunit, representing the key binding site on the heterotrimeric G protein. Inactive rhodopsin consists of the apoprotein opsin in its inactive conformation and the inverse agonist 11-*cis*-retinal (shown as red spheres), which is covalently bound to Lys 296 (shown as black spheres) via a Schiff base. In Meta II, the agonist all-

trans-retinal (shown as blue spheres) is covalently linked to Lys 296. The G-protein-interacting conformation of Meta II (right panel) is identified by the bound Gα fragment (shown in purple). Transmembrane helices (numbered 1–7) are followed by a cytoplasmic helix H8. The Meta II models lack 22 C-terminal residues, the C terminus of rhodopsin is not shown. The extracellular domain contains four β-strands (coloured in cyan). Glycosylation at Asn 2 and Asn 15 and palmitoylation at Cys 322 and Cys 323 (at the end of H8) are not shown.

binding site for all-*trans*-retinal appears to be preformed in Ops* because the presence of retinal causes only a small adjustment of side chains (Trp 265^{6,48}, Tyr 268^{6,51}, Met 207^{5,42}, His 211^{5,46} and Glu 122^{3,37}) while the Lys 296^{7,43} side chain, which is more flexible in ligand-free Ops* (ref. 6), becomes ordered due to its linkage with retinal (Supplementary Fig. 6). The Meta II-GαCT2 structure reveals a potential

hydrogen bonding network in which a water molecule links the Schiff base nitrogen with the side chains of Ser 186^{EL2} and Glu 181^{EL2}, the latter hydrogen-bonded to the side chain of Tyr 268^{6,51} (Fig. 2 and Supplementary Fig. 7). A further water molecule constrains the main chain carbonyl of Glu 181^{EL2}, the side chain of Ser 186^{EL2} and the NH-groups of Ser 186^{EL2} and Cys 185^{EL2}. From Cys 185^{EL2} the hydrogen bonding

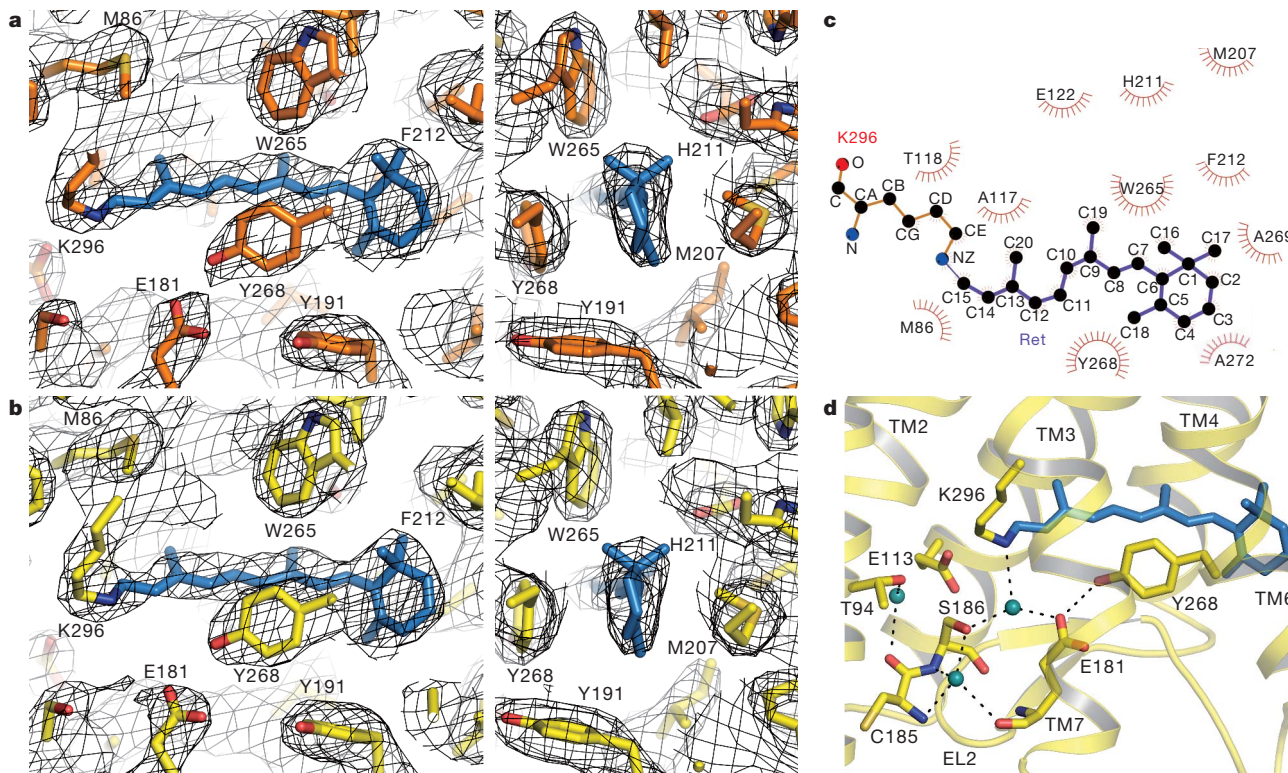


Figure 2 | Retinal binding pocket of Meta II. **a, b**, Retinal and neighbouring residues shown as stick model for Meta II (orange) and Meta II-GαCT2 (yellow), respectively, with $2F_o - F_c$ electron density map contoured at 1.0σ (grey mesh). View from within the membrane, right panel rotated 90° with respect to the left. Retinal is shown as blue stick model. **c**, Hydrophobic

interactions between all-*trans*-retinal and residues in the ligand binding pocket of the opsin protein moiety. Residues with closest distances less than 4 Å are considered to be in van der Waals contact. **d**, Potential hydrogen bonding interactions around retinal in Meta II-GαCT2. Water molecules are shown as blue spheres.

network extends via a water molecule to the side chain of Thr 94^{2,61}. In the inactive rhodopsin conformation, a different network is found in which the Schiff base nitrogen is within hydrogen-bonding distance to the side chain of Glu 113^{3,28} from which the network can extend to Thr 94^{2,61} and via Ser 186^{EL2} to Glu 181^{EL2} (refs 9, 10). Together Glu 113^{3,28} and Glu 181^{EL2} form the counterion for the protonated Schiff base in inactive rhodopsin. From Fourier transform infrared (FTIR) studies on rhodopsin activation it was predicted that the main contribution to this complex counterion changes from Glu 113^{3,28} to Glu 181^{EL2} during the transition to Meta I (ref. 12). This is also reflected in the Meta II structure. Relative to rhodopsin, the distance between the Schiff base and the side-chain carboxyl group of Glu 113^{3,28} increases from 3.5 to 5.3 Å, whereas for Glu 181^{EL2} the corresponding distance decreases from 7.1 to 5.0 Å.

The Meta II-G α CT2 structure reveals additional structural water molecules next to amino acids in conserved motifs, extending the hydrogen-bonding network from Asn 302^{7,49} of the NPxxY(x)_{5,6}F motif to Glu 134^{3,49} of the E(D)RY motif. The lack of an interaction between Glu 134^{3,49} and Arg 135^{3,50} in Meta II-G α CT2 enables Arg 135^{3,50} to bind G α CT2 by forming a hydrogen bond to the backbone carbonyl oxygen of Val 347 of G α CT2. TM2 and TM7 are linked via a water molecule between Ser 298^{7,45} and highly conserved residues Asp 83^{2,50} and Asn 302^{7,49}. Structural water molecules are conserved in GPCRs¹³ and changes in radiolytic hydroxyl radical labelling have indicated activation-induced alterations in the local structures in which they are embedded¹⁴. Our data provide further insight into how water molecules are rearranged in the course of receptor activation (Supplementary Fig. 7).

In contrast to rhodopsin^{8–10}, but like Ops* (refs 6, 15), the Meta II structures show two openings of the retinal binding pocket (Supplementary Fig. 8). The opening between TM1 and TM7 is caused by a rotamer change of Phe 293^{7,40} relative to rhodopsin, which occurs in the course of a rotational rearrangement of the extracellular part of TM7 below the retinal attachment site Lys 296^{7,43} (Supplementary Fig. 9). The opening between TM5 and TM6 arises mainly from side-chain changes of Phe 208^{5,43}, Phe 212^{5,47} and Phe 273^{6,56} which are also observed in the Ops* structure⁶. The location of retinal in its binding site fulfils most distance constraints obtained from dipolar-assisted rotational resonance (DARR) NMR spectroscopy (Supplementary Table 2)^{16,17}. A few larger distance deviations indicate that perhaps different Meta II sub-states were observed in the NMR and X-ray experiments (see Scheme 1 in the Supporting Information).

Meta II shows, like Ops*, two gross conformational changes with respect to rhodopsin (Supplementary Fig. 3)^{6,7}. Motion of TM5 towards TM6 is linked to a rearrangement of the TM3-TM5 hydrogen bonding network between Glu 122^{3,37}, Trp 126^{3,41} and His 211^{5,46}. In Meta II, the side chains of Glu 122^{3,37} and His 211^{5,46} are close enough to form a hydrogen bond (Supplementary Figs 6 and 7). The second extensive conformational change involves TM6, which undergoes a rotational tilt resulting in a motion of its cytoplasmic end away from TM3 (Figs 3 and 4). For the corresponding structural rearrangement of residues in the conserved motifs E(D)RY (in TM3) and NPxxY(x)_{5,6}F (in TM7-H8) see the Supplementary Discussion.

Unexpectedly, all-*trans*-retinal in the reversibly formed Meta II appears to be rotated around its long axis compared with the dark state^{8–10}, and the early photoproducts Batho¹⁸ and Lumi¹⁹ (Fig. 3). The significant electron density observed for the methyl groups of the retinal polyene chain and β -ionone ring gives a preference for modelling retinal rotated approximately upside-down relative to Lumi (Fig. 3a) and non-satisfactory results when only the polyene chain was rotated but the β -ionone ring was kept as in Lumi (6-*s-trans* conformation). Retinal rotation is unlikely to occur with the packing seen in the Meta II crystal (Fig. 2c), but small TM movements in the receptor activation process, which can occur as early as in Meta I²⁰, may well provide the necessary space. The Meta II structure does neither reflect the proposed ‘toggle switch’ rotational change of Trp 265^{6,48}

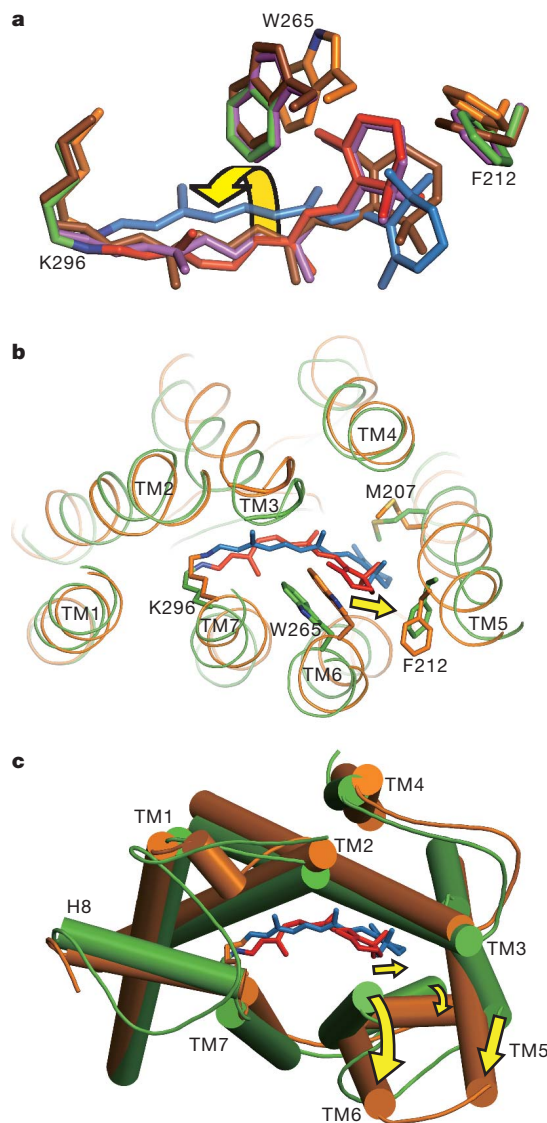


Figure 3 | Superposition of rhodopsin, Batho, Lumi and Meta II. **a**, View from within the membrane on superposition of retinal in rhodopsin (PDB 1U19), Batho (PDB 2G87), Lumi (PDB 2HPY) and Meta II. Amino acid side chains and retinal are shown as stick models. Protein/retinal are coloured green/red (rhodopsin), orange/blue (Meta II), purple/purple (Batho) and brown/brown (Lumi), respectively. The arrow indicates rotation of retinal from Lumi to Meta II. **b**, **c**, Superposition of rhodopsin and Meta II viewed from the cytoplasmic surface. Helices are shown as cartoon or cylinder models, retinal and amino acid side chains as stick models. For clarity, cytoplasmic ends of helices are clipped in **b**. Colours as in **a**. Elongation of retinal and interaction of the β -ionone-ring moiety with TM5/TM6 induces a rigid body-type rotational tilt of TM6 and motion of TM5 (indicated by arrows).

side chain^{17,21–23} nor the displacement of loop EL2 from the retinal binding site¹⁶. An explanation would be that these structural changes are transient and involved in switching to rather than maintaining the active conformation (see Supplementary Discussion and Supplementary Fig. 3). The Meta II structure also shows a positively twisted 6-*s-cis* conformation regarding the bond between carbon atoms C6 and C7 where retinal's β -ionone ring and polyene chain are connected. A negatively twisted 6-*s-cis* conformation as in the rhodopsin dark state, Batho and Lumi, was predicted for Meta I and Meta II by NMR spectroscopy^{17,24–26}, although flexibility of retinal even in the inactive receptor

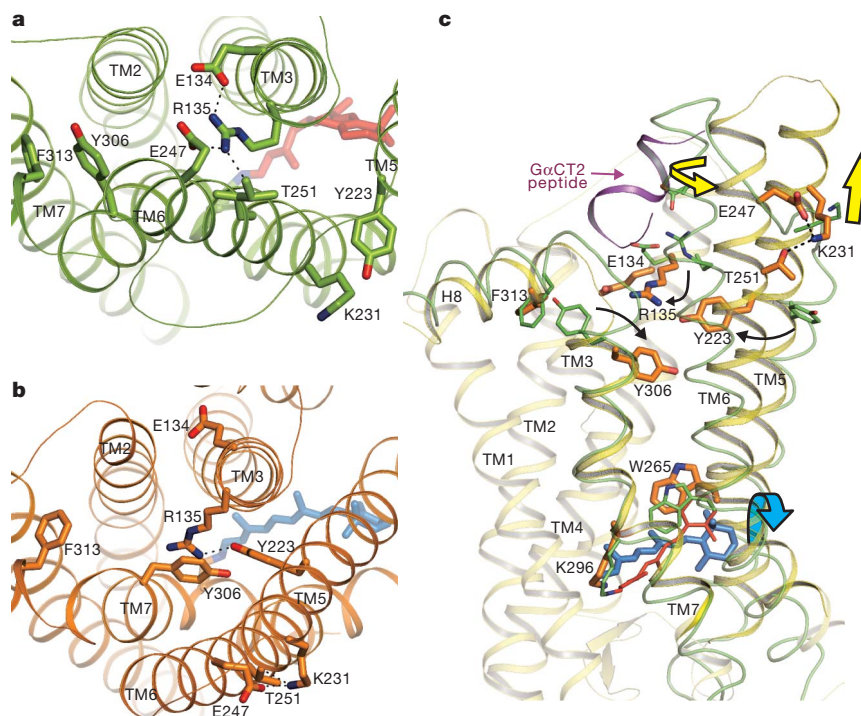


Figure 4 | Conserved E(D)RY and NPxxY(x)_{5,6}F regions. **a**, Rhodopsin and **b**, Meta II viewed from the cytoplasmic surface. Residues in the conserved E(D)RY and NPxxY(x)_{5,6}F motifs in TM3 and TM7/H8, respectively, constrain rhodopsin (green) in the inactive state. A hydrogen bonded network termed ‘ionic lock’ is formed between TM3 and TM6 including Glu 134^{3,49}, Arg 135^{3,50}, Glu 247^{6,30} and Thr 251^{6,34}. Tyr 306^{7,53} and Phe 313^{7,60} show an electrostatic interaction and are part of a functional microdomain between TM7 and H8. In Meta II and Meta II–GαCT2, the conserved residues form a different set of interactions. Glu 247^{6,30}–Lys 231^{5,66}–Thr 251^{6,34} and Tyr 223^{5,58}–Arg 135^{3,50}

state may allow negatively and positively twisted 6-*s-cis* conformations²⁷. An artificial pigment, in which ring rotation about the C6–C7 bond is inhibited, was reported to show lower but still substantial activity towards the G protein²⁸. The positive C5–C6–C7–C8 dihedral angle may thus be preferred but not mandatory to form an active state. Generally, a special status of the retinal in the active Meta II form fits well to previous FTIR spectroscopic results, which have shown that photo-induced back reactions are different between active Meta II and its inactive precursors (Supplementary Discussion).

The X-ray analysis of the reversibly formed Meta II has shown that the ligand fits readily into the binding pocket of Ops* and displays the same relaxed conformation as in pure retinal crystals. Ops* remains in its overall structure almost unchanged when the all-*trans*-retinal ligand binds, and the cytoplasmic G protein interaction domain of Ops* and Meta II carry the same insignia of an active state. On the other hand, comparison of rhodopsin and Meta II structures indicates that the retinal must go through complex elongation and torsional motions of its polyene chain and of the β-ionone ring. The motions of the ligand are likely to correspond to identified intermediates in the light induced pathway. Further structural and spectroscopic analyses may teach us how retinal analogues are less efficiently interrogated by the protein (ref. 29 and references therein), and may open insights into partial agonism and β-arrestin biased ligands at GPCRs. A more general lesson to be learned from this work on rhodopsin is that for reaching the active GPCR state, the receptor can make use of the conformational flexibility of the ligand and the variability of its interaction with the binding site.

METHODS SUMMARY

Opsin was extracted from bovine rod outer segment disc membranes using *n*-octyl-β-D-glucopyranoside and *n*-dodecyl-β-D-maltopyranoside. Solubilized

opsin was crystallized alone or in the presence of GαCT2 peptide (NH₂-ILENLKDVGLF-COOH; Gα_{340–350}(K341L,C347V)) by hanging-drop vapour diffusion in a mixture of ammonium sulphate and sodium acetate buffer at pH 5.0–5.8. Hundreds of opsin and opsin–GαCT2 crystals grown within 5 days were systematically soaked with all-*trans*-retinal for up to 30 h in the dark to reconstitute Meta II and Meta II–GαCT2, respectively. Crystals were cryoprotected in 10% trehalose and frozen in liquid nitrogen for screening and X-ray analysis at the synchrotron BESSY II, Berlin. The Meta II structures were solved by molecular replacement using opsin (PDB accession 3DQB) as search model.

The superscript denotes N terminus (NT; for example, Asn 2^{NT}), Helix 8 (H8; for example, Phe 313^{H8}), extracellular loop (EL; for example, Glu 181^{E12}), and GPCR numbering according to Ballesteros and Weinstein³⁰. The most conserved residue in a helix among rhodopsin-like GPCRs is designated x.50, where x is the helix number. All other residues on the same helix are numbered relative to this reference residue. Based on extensive sequence alignment, Smith grouped residues into signature-conserved, group-conserved or subfamily-specific¹⁷. Underlined superscript in GPCR numbering indicates GPCR signature residues (most conserved residues; Asn 55^{1.50}, Asp 83^{2.50}, Glu 134^{3.49}, Arg 135^{3.50}, Trp 161^{4.50}, Pro 215^{5.50}, Tyr 223^{5.58}, Trp 265^{6.48}, Pro 267^{6.50}, Asn 302^{7.49}, Pro 303^{7.50}, Tyr 306^{7.53}). Note that in ref. 17, based on a larger GPCR database, Leu 79 and Tyr 223 were attributed the most conserved residues on TM2 and TM5, respectively.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 September 2010; accepted 5 January 2011.

Published online 9 March 2011.

- Rosenbaum, D. M., Rasmussen, S. G. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **459**, 356–363 (2009).
- Palczewski, K. G protein-coupled receptor rhodopsin. *Annu. Rev. Biochem.* **75**, 743–767 (2006).
- Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl Acad. Sci. USA* **105**, 7439–7444 (2008).

4. Hofmann, K. P. *et al.* A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem. Sci.* **34**, 540–552 (2009).
5. Choe, H.-W., Park, J. H., Kim, Y. J. & Ernst, O. P. Transmembrane signaling by GPCRs: insight from rhodopsin and opsin structures. *Neuropharmacology* **60**, 52–57 (2011).
6. Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H.-W. & Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454**, 183–187 (2008).
7. Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497–502 (2008).
8. Palczewski, K. *et al.* Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **289**, 739–745 (2000).
9. Okada, T. *et al.* The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **342**, 571–583 (2004).
10. Li, J., Edwards, P. C., Burghammer, M., Villa, C. & Schertler, G. F. Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **343**, 1409–1438 (2004).
11. Siebert, F. Application of FTIR spectroscopy to the investigation of dark structures and photoreactions of visual pigments. *Isr. J. Chem.* **35**, 309–323 (1995).
12. Lücke, S. *et al.* The role of Glu181 in the photoactivation of rhodopsin. *J. Mol. Biol.* **353**, 345–356 (2005).
13. Angel, T. E., Chance, M. R. & Palczewski, K. Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors. *Proc. Natl Acad. Sci. USA* **106**, 8555–8560 (2009).
14. Angel, T. E., Gupta, S., Jastrzebska, B., Palczewski, K. & Chance, M. R. Structural waters define a functional channel mediating activation of the GPCR, rhodopsin. *Proc. Natl Acad. Sci. USA* **106**, 14367–14372 (2009).
15. Hildebrand, P. W. *et al.* A ligand channel through the G protein coupled receptor opsin. *PLoS ONE* **4**, e4382 (2009).
16. Ahuja, S. *et al.* Helix movement is coupled to displacement of the second extracellular loop in rhodopsin activation. *Nature Struct. Mol. Biol.* **16**, 168–175 (2009).
17. Smith, S. O. Structure and activation of the visual pigment rhodopsin. *Annu. Rev. Biophys.* **39**, 309–328 (2010).
18. Nakamichi, H. & Okada, T. Crystallographic analysis of primary visual photochemistry. *Angew. Chem. Int. Edn Engl.* **45**, 4270–4273 (2006).
19. Nakamichi, H. & Okada, T. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc. Natl Acad. Sci. USA* **103**, 12729–12734 (2006).
20. Ye, S. *et al.* Tracking G-protein-coupled receptor activation using genetically encoded infrared probes. *Nature* **464**, 1386–1389 (2010).
21. Shi, L. *et al.* $\beta 2$ adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *J. Biol. Chem.* **277**, 40989–40996 (2002).
22. Crocker, E. *et al.* Location of Trp265 in metarhodopsin II: implications for the activation mechanism of the visual receptor rhodopsin. *J. Mol. Biol.* **357**, 163–172 (2006).
23. Nygaard, R., Frimurer, T. M., Holst, B., Rosenkilde, M. M. & Schwartz, T. W. Ligand binding and micro-switches in 7TM receptor structures. *Trends Pharmacol. Sci.* **30**, 249–259 (2009).
24. Salgado, G. F. *et al.* Solid-state ^2H NMR structure of retinal in metarhodopsin I. *J. Am. Chem. Soc.* **128**, 11067–11071 (2006).
25. Ahuja, S. *et al.* 6-*s-cis* conformation and polar binding pocket of the retinal chromophore in the photoactivated state of rhodopsin. *J. Am. Chem. Soc.* **131**, 15160–15169 (2009).
26. Brown, M. F., Salgado, G. F. & Struts, A. V. Retinal dynamics during light activation of rhodopsin revealed by solid-state NMR spectroscopy. *Biochim. Biophys. Acta* **1798**, 177–193 (2010).
27. Lau, P. W., Grossfield, A., Feller, S. E., Pitman, M. C. & Brown, M. F. Dynamic structure of retinylidene ligand of rhodopsin probed by molecular simulations. *J. Mol. Biol.* **372**, 906–917 (2007).
28. Fujimoto, Y. *et al.* On the bioactive conformation of the rhodopsin chromophore: absolute sense of twist around the 6-*s-cis* bond. *Chem. Eur. J.* **7**, 4198–4204 (2001).
29. Knierim, B., Hofmann, K. P., Gartner, W., Hubbell, W. L. & Ernst, O. P. Rhodopsin and 9-demethyl-retinal analog: effect of a partial agonist on displacement of transmembrane helix 6 in class A G protein-coupled receptors. *J. Biol. Chem.* **283**, 4967–4974 (2008).
30. Ballesteros, J. A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G-protein coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Engelmann, C. Koch and B. Bauer for technical assistance, and F. Siebert and W. Hubbell for critically reading the manuscript. We are grateful to the European Synchrotron Radiation Facility (ESRF, Grenoble), D. von Stetten and A. Royant of the ID29S-Cryobench (ESRF, Grenoble) and U. Müller and the scientific staff of the BESSY-MX/Helmholtz Zentrum Berlin für Materialien und Energie at beamlines BL 14.1 and BL 14.2, where the data were collected, for continuous support. This work was supported by the DFG Sfb449 (to O.P.E.), Sfb740 (to O.P.E. and K.P.H.) and an Advanced Investigator ERC grant (to K.P.H.) and by the Canada Research Chairs Program (to E.F.P.). H.-W.C. gratefully acknowledges the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0002738) and CBNU funds for overseas research 2009. Y.J.K. thanks the Leibniz Graduate School of Molecular Biophysics, Berlin, for a scholarship.

Author Contributions H.-W.C., Y.J.K. and J.H.P. are joint first authors. H.-W.C., Y.J.K., J.H.P. performed preparation and crystallization of opsin/opsin–G α CT2. H.-W.C. performed the soaking experiment of both crystals. O.P.E. designed G α CT2. H.-W.C., Y.J.K., J.H.P., P.S., O.P.E. performed the data collection. Y.J.K., P.S., N.K. performed the structural analysis of Meta II, and J.H.P., P.S., E.F.P. performed the structural analysis of Meta II•G α CT2. T.M. performed the spectroscopic and biochemical analysis. H.-W.C., N.K., K.P.H., P.S., O.P.E. analysed data and H.-W.C., K.P.H., O.P.E. wrote the paper with contributions from all authors.

Author Information Atomic coordinates and structure factors for the reported structure have been deposited in the Protein Data Bank with the accession codes 3PQR and 3PXO. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.-W.C. (hwchoe@bnu.ac.kr), K.P.H. (klaus_peter.hofmann@charite.de) or O.P.E. (oliver.ernst@utoronto.ca).

METHODS

Crystallization. Opsin was prepared and crystallized by hanging drop vapour diffusion at 277 K essentially as described^{6,7}. The solubilization buffer contained a mixture of 1% *n*-octyl- β -D-glucopyranoside and 0.02% *n*-dodecyl- β -D-maltopyranoside or *n*-octyl- β -D-glucopyranoside alone. The precipitant solution was 3.0–3.4 M $(\text{NH}_4)_2\text{SO}_4$ in 0.1 M sodium acetate buffer (pH 5.0–5.8). A synthetic peptide was used for co-crystallization of opsin with a C-terminal fragment derived from the α -subunit of the G protein transducin. The peptide G α CT2 (sequence NH₂-ILENLKDVGLF-COOH) is a derivative of the high-affinity G α CT peptide³¹ and has a similar affinity. Crystals appeared within 2 days and grew further for 3 days to reach dimensions 0.2 mm \times 0.3 mm \times 0.3 mm. All-*trans*-retinal was purchased from Sigma and further purified by normal phase HPLC using a LiChrosorb Si 60 (5 μ m) 300 mm \times 20 mm column with 5% diethylether in heptane. Opsin crystals were soaked in the dark with all-*trans*-retinal for up to 30 h to reconstitute Meta II. Crystals were then transferred under orange light to cryoprotectant consisting of 10% trehalose in crystallization buffer and immediately frozen in liquid nitrogen.

Structure analysis. Diffraction data collection was performed at 100 K using synchrotron X-ray sources at BESSY II, Berlin, Germany, and ESRF, Grenoble, France. Best diffraction data were collected at beamline BL 14.2 at BESSY II, at $\lambda = 0.91842$ Å. The crystal to MX-225 CCD detector distance was fixed at 230 mm for Meta II and 260 mm for Meta II–G α CT2, respectively. The rotation increment for each frame was 0.5° with an exposure time of 6 s (Meta II) or 12 s (Meta II–G α CT2). All images were indexed, integrated and scaled using the XDS program package³² and CCP4 program SCALA³³. Meta II and Meta II–G α CT2 crystals belong to rhombohedral space group *H*32 ($a = 241.82$ Å, $b = 241.82$ Å, $c = 111.80$ Å, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$ and $a = 241.55$ Å, $b = 241.55$ Å, $c = 109.87$ Å, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$, respectively). Supplementary Table 1 summarizes the statistics for crystallographic data collection and structural refinement.

Initial phases for Meta II and Meta II–G α CT2 were obtained by conventional molecular replacement protocol (rotation, translation, rigid body fitting) using opsin from the opsin–G α CT structure (PDB accession 3DQB) as initial search model. Molecular replacement was achieved using the CCP4 program PHASER³³ by first placing the opsin monomer (rotation function (RFZ): $Z = 14.5$; translation function (TFZ): $Z = 53.2$ for Meta II and rotation function (RFZ): $Z = 11.7$; translation function (TFZ): $Z = 63.5$ for Meta II–G α CT2, respectively; RFZ and TFZ as defined by PHASER). In subsequent steps, torsion angle molecular dynamics, simulated annealing using a slow-cooling protocol and a maximum likelihood target function, energy minimization, and *B*-factor refinement by the program CNS³⁴ were carried out in the resolution range 33.5–3.0 Å (Meta II) and 34.9–2.85 Å (Meta II–G α CT2). After the first round of refinement, the retinal chromophore was clearly visible in the electron density of both σ_A -weighted $2F_o - F_c$ maps, as well as in the σ_A -weighted simulated annealing omit density maps (Supplementary Fig. 4). Restrained, individual *B*-factors were refined and the crystal structure was finalized by the CCP4 program REFMAC5 and CCP4 (ref. 33). The final models have agreement factors R_{free} and R_{cryst} of 24.5% and 21.6% (Meta II), and 25.0% and 21.7% (Meta II–G α CT2), respectively. Manual rebuilding of the Meta II models and electron density interpretation were performed after each refinement cycle using the program COOT³⁵. Structure validation was performed with the programs PROCHECK³⁶ and WHAT_CHECK³⁷. Potential hydrogen bonds and van der Waals contacts were analysed using the programs HBPLUS³⁸ and LIGPLOT³⁹. All crystal structure superpositions of backbone alpha

carbon traces were performed using CCP4 program LSQKAB³³. All molecular graphics representations were created using PyMol⁴⁰.

Ultraviolet-visible spectroscopy. Ultraviolet-visible spectroscopy of Meta II was performed using a NanoDrop 1000 Spectrophotometer (Thermo Scientific) at room temperature and a CARY 50 Ultraviolet-visible Spectrophotometer (Varian) at 20 °C. Meta II crystals were washed thoroughly three times in crystallization buffer to reduce excess all-*trans*-retinal and were measured immediately in the same buffer. In some cases, Meta II crystals were dissolved in 1% *n*-octyl- β -D-glucopyranoside. Spectra were obtained before and after acid denaturation by addition of 1/10 volume 100 mM sulphuric acid.

Transducin activation assay. As a monitor for Gt activation, changes in intrinsic fluorescence intensity of the G α -subunit upon exchange of GDP to GTP γ S were quantified^{41–43}. All measurements were carried out using a SPEX fluorolog II spectrofluorometer equipped with a 450 W xenon arc lamp. For all activation measurements, settings were $\lambda_{\text{ex}} = 300$ nm and $\lambda_{\text{em}} = 345$ nm with an integration time of 1 s. Gt activation rates were measured with 2.5 nM opsin, 6 μ M GTP γ S, 2 mM DTT, and 0.006% (w/v) *n*-dodecyl- β -D-maltopyranoside in a final volume of 650 μ l (10 mm \times 4 mm cuvette with stirring bar). All samples were equilibrated at 20 °C for 4 min while being irradiated with yellow light (Schott GG 495 long-pass filter). Then reactions were triggered by addition of GTP γ S after recording basic fluorescence levels for 50 s. After recording the fluorescence change at the initial opsin concentration, activation of the whole Gt pool was achieved by adding 10 nM purified rhodopsin to the reaction. The concentration of intact opsin in every sample was calculated from the quantity of reconstituted rhodopsin by 500 nm absorption after addition of excess 11-*cis*-retinal to an aliquot of the sample measured. The 500 nm rhodopsin absorption was determined from difference spectra dark minus light-activated rhodopsin.

31. Herrmann, R. *et al.* Sequence of interactions in receptor-G protein coupling. *J. Biol. Chem.* **279**, 24283–24290 (2004).
32. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
33. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
34. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
35. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
36. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993).
37. Hoof, R. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* **381**, 272 (1996).
38. McDonald, I. K. & Thornton, J. M. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238**, 777–793 (1994).
39. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–134 (1995).
40. DeLano, W. L. The PyMOL Molecular Graphics System. (DeLano Scientific, San Carlos, California, USA, 2002).
41. Fahmy, K. & Sakmar, T. P. Regulation of the rhodopsin-transducin interaction by a highly conserved carboxylic acid group. *Biochemistry* **32**, 7229–7236 (1993).
42. Ernst, O. P., Bieri, C., Vogel, H. & Hofmann, K. P. Intrinsic biophysical monitors of transducin activation: fluorescence, UV-visible spectroscopy, light scattering, and evanescent field techniques. *Methods Enzymol.* **315**, 471–489 (2000).
43. Ernst, O. P., Gramse, V., Kolbe, M., Hofmann, K. P. & Heck, M. Monomeric G protein-coupled receptor rhodopsin in solution activates its G protein transducin at the diffusion limit. *Proc. Natl Acad. Sci. USA* **104**, 10859–10864 (2007).

The structural basis of agonist-induced activation in constitutively active rhodopsin

Jörg Standfuss^{1,2}, Patricia C. Edwards², Aaron D'Antona³, Maikel Fransen¹, Guifu Xie³, Daniel D. Oprian³ & Gebhard F. X. Schertler^{1,2}

G-protein-coupled receptors (GPCRs) comprise the largest family of membrane proteins in the human genome and mediate cellular responses to an extensive array of hormones, neurotransmitters and sensory stimuli. Although some crystal structures have been determined for GPCRs, most are for modified forms, showing little basal activity, and are bound to inverse agonists or antagonists. Consequently, these structures correspond to receptors in their inactive states. The visual pigment rhodopsin is the only GPCR for which structures exist that are thought to be in the active state^{1,2}. However, these structures are for the apoprotein, or opsin, form that does not contain the agonist all-*trans* retinal. Here we present a crystal structure at a resolution of 3 Å for the constitutively active rhodopsin mutant Glu 113 Gln^{3–5} in complex with a peptide derived from the carboxy terminus of the α -subunit of the G protein transducin. The protein is in an active conformation that retains retinal in the binding pocket after photoactivation. Comparison with the structure of ground-state rhodopsin⁶ suggests how translocation of the retinal β -ionone ring leads to a rotation of transmembrane helix 6, which is the critical conformational change on activation⁷. A key feature of this conformational change is a reorganization of water-mediated hydrogen-bond networks between the retinal-binding pocket and three of the most conserved GPCR sequence motifs. We thus show how an agonist ligand can activate its GPCR.

In the dark, rhodopsin contains a covalently bound 11-*cis* retinal chromophore that preferentially binds to the inactive state and therefore functions as an inverse agonist in the visual system. On exposure to light, the retinal isomerizes to the all-*trans* form, initiating a series of conformational changes leading to the transient signalling state metarhodopsin II.

Crystal structures of bovine rhodopsin have been determined for the inactive state of the native protein bound to 11-*cis* retinal and from crystals of the protein following photoisomerization of the chromophore to the all-*trans* form⁸. However, these structures represent early photointermediates in the activation pathway^{9–11} that have not yet undergone the critical conformational alteration required for activation of the G protein.

Recently, two similar structures have been determined for the apoprotein form opsin^{1,2}. These are thought to represent activated forms of the receptor because one was determined for opsin bound to a peptide derived from the C-terminal tail of the α -subunit of the G protein transducin (the G α CT peptide) known to bind preferentially and stabilize the active intermediate metarhodopsin II¹². Although the opsin structures have been of immense value, a complete understanding of the active state requires structures in which the all-*trans* retinal agonist is included in the ligand-binding pocket of the receptor.

We have obtained such a structure by crystallization of the Glu 113 Gln^{3,28} mutant (superscripts denote Ballesteros–Weinstein general GPCR numbering¹³) that drastically slows hydrolysis and dissociation of all-*trans* retinal from the photoactivated protein by

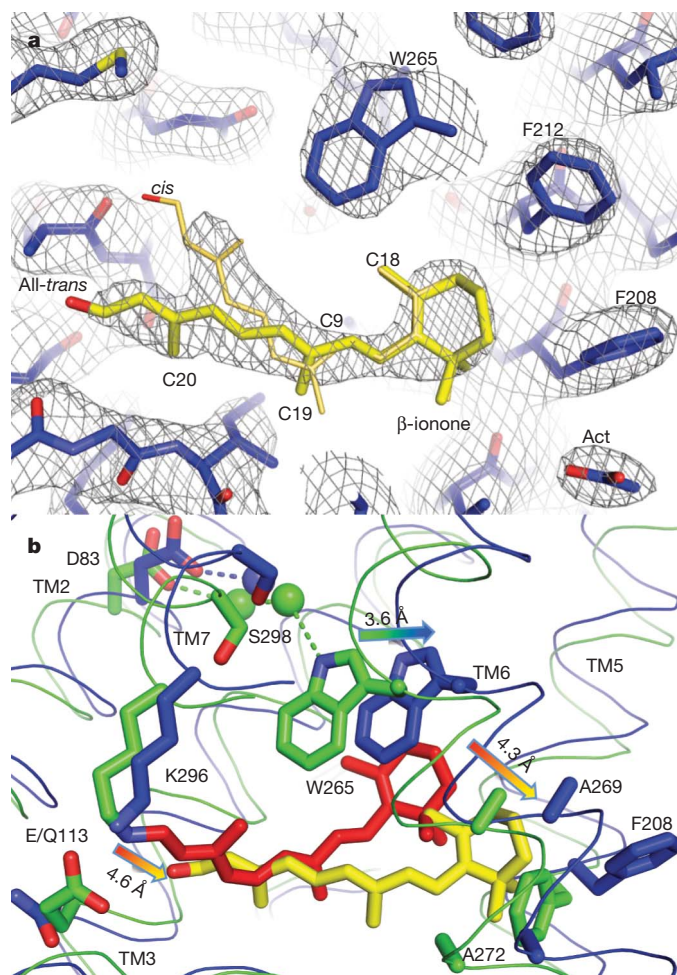


Figure 1 | Conformational changes in the retinal binding pocket.

a, $2F_o - F_c$ map (contoured at 1.5σ) of the retinal-binding pocket. The retinal β -ionone ring is well resolved whereas the density of the polyene chain broadens towards the end facing Lys 296^{7,43}. Occupancy refinements indicated a 60:40 mixture of all-*trans* retinal and a mixed population. An acetate molecule (Act) is packed between Phe 208^{5,43} and Phe 276^{6,59}, blocking a proposed retinal entry/exit channel¹. **b**, Superposition of Glu 113 Gln/G α CT (blue; PDB ID, 2X72) with ground-state rhodopsin (green; PDB ID, 1GZM⁶). The β -ionone ring of all-*trans* retinal (yellow) is shifted 4.3 Å towards the left between TM5 and TM6. This releases Trp 265^{6,48} from its locked position in the ground state and disrupts a water-mediated TM6–TM7 interaction involving Ser 298^{7,45} (ref. 28). The salt bridge between Glu 113^{3,28}, as counterion to the protonated Schiff base, and Lys 296^{7,43} is broken in the Glu 113 Gln/G α CT structure, removing the restraining TM3 and TM7 in the ground state⁵.

¹Paul Scherrer Institut, 5232 Villigen PSI, Switzerland. ²MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK. ³Department of Biochemistry and Volen Center for Complex Systems, Brandeis University, Waltham, Massachusetts 02454, USA.

neutralization of the retinal counterion⁴. Furthermore, the mutant shifts the classical metarhodopsin I/metarhodopsin II equilibrium towards active metarhodopsin II and maintains metarhodopsin III in an active conformation¹⁴. Finally, the opsin form of the Glu 113 Gln^{3,28} mutant is constitutively active⁵, and addition of all-*trans* retinal to the apoprotein can activate the mutant to levels comparable to those of light-activated wild-type rhodopsin^{4,15}, essentially turning rhodopsin into a GPCR that is activated by diffusible agonists. However, the Glu 113 Gln^{3,28} mutant is also significantly less stable than wild-type rhodopsin (Supplementary Fig. 1), and for this reason we used the Glu 113 Gln mutation in the context of another rhodopsin mutant, Asn 2 Cys/Asp 282 Cys, which contains an engineered disulphide bond known to enhance thermal stability of the protein without affecting activity^{14,16} or structure¹⁷.

The Glu 113 Gln/Asn 2 Cys/Asp 282 Cys triple mutant (henceforth referred to simply as the Glu 113 Gln^{3,28} mutant) was reconstituted with 11-*cis* retinal during purification and, just before crystallization, was activated by selective illumination of protonated retinal in presence of the G α CT peptide. Crystals used for data collection had a faint yellow colour (Supplementary Fig. 2) indicative of bound retinal.

The structure contains residues 1–326 of the mutant opsin and all eleven residues of the G α CT peptide, including the Lys 341 \rightarrow Leu mutation introduced to increase affinity for the receptor¹⁸. In addition, clear electron density is observed for two partially ordered lipid molecules, one molecule of octylglucoside, several water molecules and, most strikingly, retinal within the ligand-binding pocket of the receptor. With the exception of the missing oligosaccharyl chain at position 2, the recombinantly produced protein contains all post-translational modifications observed with the native protein including acetylation of the amino terminus, palmitoylation of Cys residues at positions 322 and 323, and glycosylation of Asn at position 15. Finally, the engineered disulphide between Cys residues at positions 2 and 282 is clearly visible in the mutant.

The structure of the Glu 113 Gln/G α CT complex deviates significantly from the ground state of rhodopsin (Protein Data Bank ID,

1GZM⁶), but has high similarity to the active-state structure of the opsin/G α CT complex (PDB ID, 3DQB²), with respective C α root mean squared deviations of 2.42 Å and 0.58 Å (Supplementary Fig. 3). The Glu 113 Gln/G α CT structure is also in excellent agreement with the results of high-resolution distance mapping using double electron–electron resonance (DEER) spectroscopy where pairs of nitroxide spin labels (in particular at positions 241 and 252) were used to quantify an outward movement of transmembrane helix 6 (TM6) by 5 Å that is critical to the activation process⁷ (Supplementary Table 1). Other residues not showing significant change in the DEER experiments also do not show significant difference from ground-state rhodopsin in the Glu 113 Gln/G α CT structure.

On the basis of constitutive activity of the Glu 113 Gln^{3,28} mutant, the presence of the G α CT peptide and the agreement with the results of DEER spectroscopy, we conclude that the structure of the Glu 113 Gln/G α CT complex reported here is in fact that of the active state. Electron density for retinal (Fig. 1) is clearly observed in the general area where 11-*cis* retinal is found in ground-state rhodopsin. However, density for the nearby side chain of Lys 296^{7,43} is weak, indicating that retinal is not bound to the protein covalently by a Schiff base, as in the ground state or metarhodopsin II. In addition, density for the β -ionone ring and most of the polyene chain is well defined but shows increasing disorder after position C9. We have modelled retinal in the all-*trans* conformation, but on the basis of occupancy refinement we estimate a 60:40 mixture of all-*trans* retinal and a mixed population. However, we cannot distinguish between a model based on a mixture of *cis* and *trans* isomers, catalysed for example by phosphatidylethanolamine lipids¹⁹ or Lys 296^{7,43}, and one based on a mixture of conformers arising from rotations about single bonds in the polyene chain or one with residual, covalently bound all-*trans* retinal contributing to the density.

Clearly, the structure represents an active conformation but is not identical to metarhodopsin II. We suspect that it probably corresponds to a trapped intermediate in which the retinal is either entering or exiting the ligand-binding pocket. It is well established that wild-type

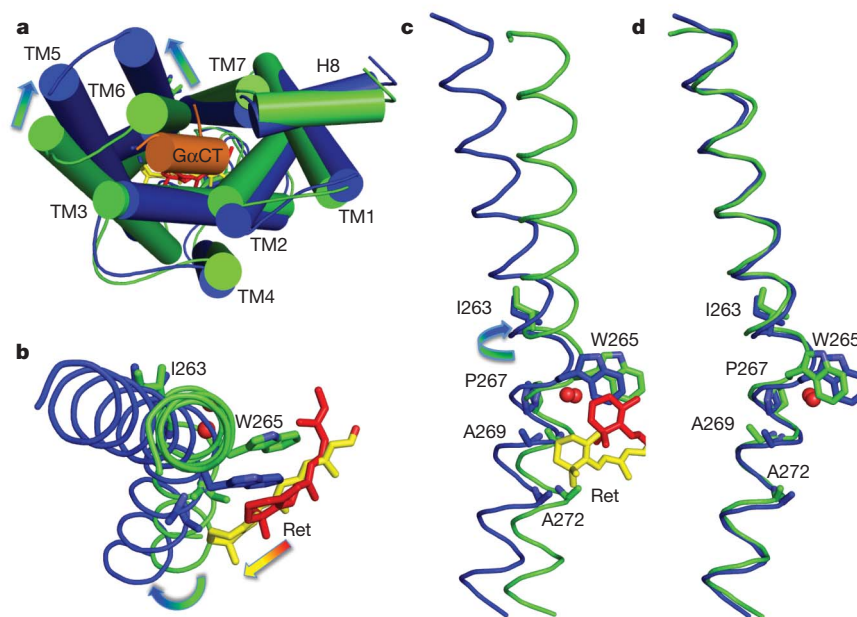
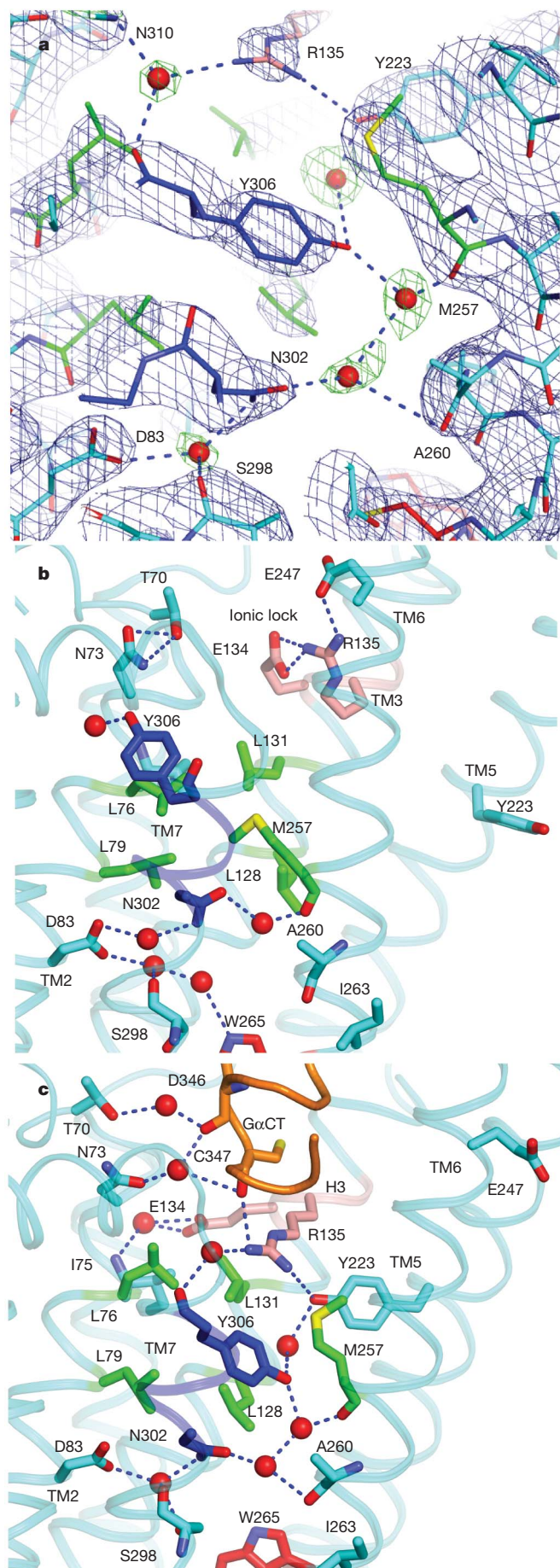


Figure 2 | Rearrangement of the heptahelix bundle and rotation of TM6.

a, Superposition of C α traces of ground-state rhodopsin (green) and Glu 113 Gln (blue) with bound G-protein peptide (orange). The main rearrangements of transmembrane helices (cylinders) that open the G-protein-binding site are indicated as arrows. The loop regions have been smoothed for clarity. **b**, **c**, Cytoplasmic-side (**b**) and membrane-side (**c**) views of TM6 (C α traces in ribbon representation) from ground-state rhodopsin and the

Glu 113 Gln/G α CT structure illustrate how 11-*cis* (red) to all-*trans* (yellow) isomerization of retinal (Ret) can release Trp 265^{6,48} and lead to rotation of TM6. Resulting conformational changes are amplified towards the cytoplasmic side by the characteristic bend of the helix introduced by Pro 267^{6,50} and water molecule 7. **d**, Superposition of TM6 alone shows that the shape of the helix is preserved during the rearrangements.



opsin transiently activates as retinal enters the binding pocket but before a covalent bond to Lys 296^{7,43} has been formed²⁰, and that both all-*trans* and 11-*cis* retinol can act as potent partial agonists²¹. In this regard, the well-ordered ionone ring end of retinal and part of the polyene chain in the Glu 113 Gln/GαCT structure is probably in the same position that it occupies in metarhodopsin II. With respect to ground-state rhodopsin, the β-ionone ring is shifted by 4.3 Å (Fig. 1) and inserted in the cleft between TM5 (residues Met 207^{5,42}, Phe 208^{5,43} and Phe 212^{5,47}) and TM6 (residues Trp 265^{6,48}, Ala 269^{6,52} and Ala 272^{6,55}). This position is in agreement with two-dimensional dipolar-assisted rotational resonance NMR experiments on the retinal position in metarhodopsin II²².

The transition from inactive to active states evident on comparison of the ground-state and active-state structures is accompanied by a global rearrangement of the helix bundle (Fig. 2) that shifts the cytoplasmic end of TM6 (and to a lesser extent TM5) away from the bundle core (TM1–4 and TM7). This shift is not achieved through a hinge movement but by a rotation of TM6 that leaves the shape of the helix intact. On the cytoplasmic side, the conformational change is amplified by the characteristic bend caused by Pro 267^{6,50}, which is the most conserved residue in TM6 among GPCRs and part of the ubiquitous CWXP motif (where X denotes any amino acid). Pro 267^{6,50} is in close contact with water molecule 7, which is hydrogen-bonded to Cys 264^{6,47}, Tyr 268^{6,51} and Pro 291^{7,38}, an arrangement similar to that observed in ground-state rhodopsin. This water is found also in structures of the β₁-adrenergic²³, β₂-adrenergic²⁴ and adenosine receptors²⁵, and probably forms an important architectural element in formation of the bend in TM6.

Trp 265^{6,48} of the CWXP motif is a highly conserved amino acid that is tightly packed against retinal in ground-state rhodopsin and has been identified as important for GPCR activation through early mutagenesis studies. Trp 265^{6,48} has a central role in the toggle-switch model for activation of GPCRs^{26,27}. Our structure places the indole group of Trp 265^{6,48} 3.6 Å away from its ground-state position as a consequence of rhodopsin activation. However, we do not observe the rotamer change that was originally proposed on the basis of computer simulations. Instead, Trp 265^{6,48} follows the β-ionone ring, maintaining contact with the C18 methyl group. We suggest that the critical movement of TM6 stems from a motion of the β-ionone ring of retinal against TM6 just beneath the CWXP motif, while Trp 265^{6,48} is simultaneously released from its locked ground-state position.

Of special interest in the Glu 113 Gln/GαCT structure is a cluster of densities that indicate the presence of structural water molecules (Fig. 3) between some of the most conserved residues in GPCRs. Although modelling of water is difficult at resolutions around 3 Å, we are confident in their position because omission of water molecules during refinement resulted in clear difference peaks, presumably due to tight hydrogen bonding typical for structural water molecules in the interior of membrane proteins. Our interpretation is further strengthened by the presence of a similar water-mediated hydrogen-bond

Figure 3 | Rearrangement of water-mediated hydrogen-bond networks.

a, Our crystallographic data indicates water molecules that mediate interactions between the retinal-binding region and the GαCT-binding site. Omission of water molecules during simulated annealing results in strong electron density difference peaks (blue mesh, $2F_o - F_c$ map contoured at 2.0σ ; green mesh, $F_o - F_c$ map contoured at 3.0σ), demonstrating a high degree of local order. **b**, The retinal-binding pocket of ground-state rhodopsin is connected with Asn 302^{7,49} of the NPXXY motif (blue) via water-mediated hydrogen-bond networks⁶. A region called the hydrophobic barrier (green) separates this network from the E(D)RY motif (pink) critical for G-protein activation. **c**, In the active Glu 113 Gln/GαCT structure, rotation of TM6 disrupts the water-mediated link between Trp 265^{6,48} and Ser 298^{7,45} in TM7 and reorganizes the ground-state hydrogen-bond network. The hydrophobic barrier opens and Tyr 306^{7,53} of the NPXXY motif (blue) in TM7 and Tyr 223^{5,58} in TM5 rearrange to fill the resulting gap to extend the hydrogen-bond network towards the E(D)RY motif (pink) and the GαCT peptide (orange).

network found in ground-state rhodopsin that begins at Trp 265^{6,48}, involves the conserved Asp 83^{2,50} in TM2 and Ser 298^{7,45} and Asn 310^{8,47} (of the NPXXY motif) in TM7, and ends just below the hydrophobic barrier of six residues between TM2, TM3 and TM6 (Leu 76^{2,43}, Leu 79^{2,46}, Leu 128^{3,43}, Leu 131^{3,46}, Met 253^{6,36} and Met 257^{6,40}), of which many are conserved in rhodopsin-class GPCRs.

The Glu 113 Gln/G α CT structure indicates a rearrangement of this water cluster through rotation of TM6. The water-mediated link between Trp 265^{6,48} in TM6 and Ser 298^{7,45} in TM7 is broken, whereas Ser 298^{7,45}, together with water molecule 16, Asn 55^{1,50}, Asp 83^{2,50} and Asn 302^{7,49}, continues to stabilize the unusual Pro kink introduced in TM7²⁸ by Pro 303^{7,50}. These reorganizations are comparatively minor, but they directly link changes in the CWXP motif in the retinal-binding pocket with the two most conserved residues in TM1 and TM2, and the NPXXY motif in TM7. On the cytoplasmic side, the rotation of TM6 opens the hydrophobic barrier, allowing Tyr 223^{5,58} and Tyr 306^{7,53} of the NPXXY motif to swing into the protein interior from their positions in the ground state. The phenolic side chains provide additional interactions with water molecules 2 and 14, extending the hydrogen-bond network through the opened hydrophobic barrier up to the E(D)RY motif at the cytoplasmic surface of TM3. Residues Glu 134^{3,49} and Arg 135^{3,50} of this motif are released from the warped conformation in the ground state, which opens the ionic lock interactions²⁹ of Glu 134^{3,49}/Arg 135^{3,50} and Glu 247^{6,30} and allows binding of the G α CT peptide in a position that is occupied by TM6 in the ground state. Thus, rotation of TM6 and displacement of Trp 265^{6,48} results in a hydrogen-bond network connecting residues from the protein interior in contact with retinal to those at the cytoplasmic surface critical to activation of the G protein. The

hydrogen-bond network is further extended towards the G α CT peptide by water molecule 13 bridging Tyr 306^{7,53} and Arg 135^{3,50} with Asn 310^{8,47} in the TM7–H8 junction (Fig. 3). Asn 310^{8,47} in turn interacts with water molecule 9, which forms hydrogen bonds to Asn 73^{2,40} of the receptor and the backbone carbonyl of Asp 346 in the peptide. Asp 346 is linked to Thr 70^{2,37} of cytoplasmic loop 1 through water molecule 8, and the side chain is linked to Asn 343 of the peptide and Arg 147^{3,62} of the receptor through water molecule 5. Structural water molecules are thus likely to have a vital role in binding of the G protein to the active receptor.

We have described how translocation of the retinal β -ionone ring can lead to the conformational changes that allow rhodopsin to bind its G protein (Fig. 4). In doing so, we show how an agonist is bound to the active state of a GPCR and how activation is accompanied by a reorganization of hydrogen-bond networks between some of the most conserved residues among GPCRs.

METHODS SUMMARY

We cloned and expressed the rhodopsin gene containing the stabilizing Asn 2 Cys/ Asp 282 Cys and Glu 113 Gln counterion mutations in HEK293S cells with restricted and homogenous *N*-glycosylation³⁰. The receptor was purified using a 1D4 immunoaffinity matrix and reconstituted with 11-*cis* retinal under dim red light as previously described¹⁷. Just before crystallization, we mixed purified rhodopsin with an equal weight of brain lipid extract, 10 μ M all-*trans* retinal and a tenfold molar excess of the G α CT/Lys 341 Leu peptide (ILENLKDCGLF). Rhodopsin was light-activated using a >515-nm long-pass filter to prevent exposure of free retinal and metarhodopsin II. The sample was mixed with an equal volume of 3.0–3.4 M ammonium sulphate, 100 mM sodium acetate, pH 4.5, and crystallized by vapour diffusion in the dark. We collected crystals under dim red light and soaked them in crystallization buffer containing 10% trehalose before freezing them. Diffraction data were collected at synchrotron X-ray sources (the Swiss Light Source, the European Synchrotron Radiation Facility and the Diamond Light Source), integrated using XDS and brought onto a common scale using SCALA. We obtained phases by molecular replacement with the opsin structures (PDB IDs, 3CAP¹ and 3DQB²) or ground-state rhodopsin (PDB ID, 1GZM⁶) as search model. The resulting solution was refined using iterative cycles of model building in COOT and refinement with PHENIX.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 March 2010; accepted 5 January 2011.

Published online 9 March 2011.

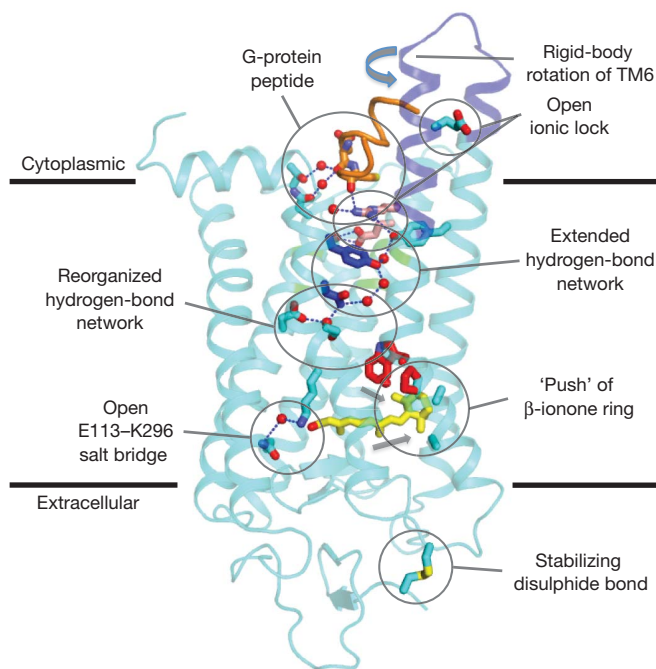


Figure 4 | Activation of rhodopsin by the agonist all-*trans* retinal. The protein backbone of Glu 113 Gln/G α CT is shown in cyan with predominant conformational changes in TM5 and TM6 (root mean squared deviation of C α atoms with respect to ground-state rhodopsin (PDB ID, 1GZM), >3.5 Å) in blue. The key regions involved in rhodopsin activation and discussed in the main text are highlighted. The side chains of the CWXP motif close to the retinal-binding site are coloured red. Side chains of the NPXXY motif (blue) extend the hydrogen-bond network through the hydrophobic barrier (green). Side chains of the E(D)RY motif, as part of the ionic lock and the G-protein-binding site, are coloured pink. The G α CT peptide is shown as an orange ribbon. The engineered disulphide bond in the extracellular domain is well isolated from the structural motifs involved in rhodopsin activation, which explains its neutral stabilizing characteristics^{14,16}.

- Park, J. H. *et al.* Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454**, 183–187 (2008).
- Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497–502 (2008).
- Zhukovsky, E. A. & Oprian, D. D. Effect of carboxylic acid side chains on the absorption maximum of visual pigments. *Science* **246**, 928–930 (1989).
- Sakmar, T. P., Franke, R. R. & Khorana, H. G. Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proc. Natl Acad. Sci. USA* **86**, 8309–8313 (1989).
- Robinson, P. R., Cohen, G. B., Zhukovsky, E. A. & Oprian, D. D. Constitutively active mutants of rhodopsin. *Neuron* **9**, 719–725 (1992).
- Li, J. *et al.* Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **343**, 1409–1438 (2004).
- Altenbach, C. *et al.* High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl Acad. Sci. USA* **105**, 7439–7444 (2008).
- Hofmann, K. P. *et al.* A G protein-coupled receptor at work: the rhodopsin model. *Trends Biochem. Sci.* **34**, 540–552 (2009).
- Nakamichi, H. & Okada, T. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc. Natl Acad. Sci. USA* **103**, 12729–12734 (2006).
- Ruprecht, J. J. *et al.* Electron crystallography reveals the structure of metarhodopsin I. *EMBO J.* **23**, 3609–3620 (2004).
- Salom, D. *et al.* Crystal structure of a photoactivated deprotonated intermediate of rhodopsin. *Proc. Natl Acad. Sci. USA* **103**, 16123–16128 (2006).
- Hamm, H. E. *et al.* Site of G protein binding to rhodopsin mapped with synthetic peptides from the alpha subunit. *Science* **241**, 832–835 (1988).
- Ballesteros, J. & Weinstein, H. Integrated methods for the construction of three dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **25**, 366–428 (1995).
- Standfuss, J., Zaitseva, E., Mahalingam, M. & Vogel, R. Structural impact of the E113Q counterion mutation on the activation and deactivation pathways of the G protein-coupled receptor rhodopsin. *J. Mol. Biol.* **380**, 145–157 (2008).

15. Cohen, G. B., Oprian, D. D. & Robinson, P. R. Mechanism of activation and inactivation of opsin: role of Glu113 and Lys296. *Biochemistry* **31**, 12592–12601 (1992).
16. Xie, G., Gross, A. K. & Oprian, D. D. An opsin mutant with increased thermal stability. *Biochemistry* **42**, 1995–2001 (2003).
17. Standfuss, J. *et al.* Crystal structure of a thermally stable rhodopsin mutant. *J. Mol. Biol.* **372**, 1179–1188 (2007).
18. Martin, E. L., Rens-Domiano, S., Schatz, P. J. & Hamm, H. E. Potent peptide analogues of a G protein receptor-binding region obtained with a combinatorial library. *J. Biol. Chem.* **271**, 361–366 (1996).
19. Groenendijk, G. W., Jacobs, C. W., Bonting, S. L. & Daemen, F. J. Dark isomerization of retinals in the presence of phosphatidylethanolamine. *Eur. J. Biochem.* **106**, 119–128 (1980).
20. Kefalov, V. J., Crouch, R. K. & Cornwall, M. C. Role of noncovalent binding of 11-*cis*-retinal to opsin in dark adaptation of rod and cone photoreceptors. *Neuron* **29**, 749–755 (2001).
21. Kono, M., Goletz, P. W. & Crouch, R. K. 11-*cis*- and all-*trans*-retinols can activate rod opsin: rational design of the visual cycle. *Biochemistry* **47**, 7567–7571 (2008).
22. Ahuja, S. *et al.* Location of the retinal chromophore in the activated state of rhodopsin. *J. Biol. Chem.* **284**, 10190–10201 (2009).
23. Warne, T. *et al.* Structure of a β_1 -adrenergic G-protein-coupled receptor. *Nature* **454**, 486–491 (2008).
24. Cherezov, V. *et al.* High-resolution crystal structure of an engineered human β_2 -adrenergic G protein-coupled receptor. *Science* **318**, 1258–1265 (2007).
25. Jaakola, V. P. *et al.* The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. *Science* **322**, 1211–1217 (2008).
26. Shi, L. *et al.* Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. *J. Biol. Chem.* **277**, 40989–40996 (2002).
27. Schwartz, T. W. *et al.* Molecular mechanism of 7TM receptor activation—a global toggle switch model. *Annu. Rev. Pharmacol. Toxicol.* **46**, 481–519 (2006).
28. Pardo, L. *et al.* The role of internal water molecules in the structure and function of the rhodopsin family of G protein-coupled receptors. *ChemBioChem* **8**, 19–24 (2007).
29. Vogel, R. *et al.* Functional role of the “ionic lock”—an interhelical hydrogen-bond network in family A heptahelical receptors. *J. Mol. Biol.* **380**, 648–655 (2008).
30. Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl Acad. Sci. USA* **99**, 13419–13424 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank X. Deupi and R. Vogel for discussions and reading of the manuscript. We thank R. Crouch for the kind gift of 11-*cis* retinal. P. J. Reeves we thank for providing the pACMVtetO vector and the HEK293S-GnTI[−] cells, and for his advice on creating stable cell lines and tetracycline-inducible expression. We also thank the staff at the macromolecular crystallography beamlines at the European Synchrotron Radiation Facility, the Diamond Light Source and the Swiss Light Source. The work was financially supported by NIH grant EY007965 (to D.D.O.), the Human Frontier Science Project programme grant RG/0052 (to D.D.O. and G.F.X.S.), the European Commission FP6 specific targeted research project LSH-2003-1.1.0-1 (to G.F.X.S.), the Marie Curie Intra European Fellowship MEIF-CT-2006-039171 (to J.S.) and the EMBO long-term fellowship ALTF 198-2006 (to J.S.).

Author Contributions The project was initiated by D.D.O. and G.F.X.S. J.S. performed cloning, initial expression and purification using essential experimental protocols and materials contributed by A.D. and D.D.O. Receptor activation and retinal binding studies were contributed by A.D. and G.X. Initial crystallization was by J.S., who also collected data and refined the structures. P.C.E. optimized expression and crystallization, performed crystal cryo-cooling and coordinated data collection. M.F. investigated the stability of mutant proteins. Manuscript preparation was performed by J.S. and D.D.O. The overall project management was by G.F.X.S.

Author Information Coordinates and structure factors have been deposited at the Protein Data Bank under accession code 2X72. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to G.F.X.S. (gebhard.schertler@psi.ch).

METHODS

Preparation of stable cell line. The rhodopsin gene containing the stabilizing Asn 2 Cys/Asp 282 Cys and Glu 113 Gln counterion mutation was cloned into the pACMVtetO vector for tetracycline-inducible expression in mammalian cells³¹ using NotI and KpnI restriction sites. HEK293S-GnTI[−] cells with restricted and homogenous *N*-glycosylation were stably transfected with this vector as previously described³². Both vector and cells were a kind gift from Philip J. Reeves, University of Essex, UK.

Large-scale expression in wave bag bioreactors. Cells were expanded as adhesion cultures in DMEM/F12 medium supplemented with FBS (10%), PenStrep (Gibco), Geneticin-G418 (200 µg ml^{−1}) and blasticidin (5 µg ml^{−1}). Cells from five fully confluent 75-cm² flasks were collected and further expanded into a 300-ml suspension culture in Freestyle Medium (Invitrogen) supplemented with FBS (5%) and PenStrep. A wave bioreactor (GE Healthcare) was used to further expand the initial suspension cultures to a volume of 9.5 l with a cell density of 2 × 10⁶ cells per millilitre. Protein expression was induced by 0.5 l medium supplemented with tetracycline and sodium butyrate for final concentrations in the wave bag of 2 µg ml^{−1} and 5 mM, respectively. Cells were collected 72 h after induction at a density of (4–5) × 10⁶ cells per millilitre. Cell pellets were washed with PBS buffer containing protease inhibitor cocktail (complete protease inhibitor cocktail tablets, Roche) and stored at −80 °C. The modelled *N*-glycan in the final structure is based on the homogenous glycosylation pattern of the HEK293-GnTI[−] cell line³² used for expression and has been built as GlcNAc₂-Man₃ with two disordered mannose sugars. Crystal contacts between two *N*-glycans and palmitoyl chains that fill the cavity between the two rhodopsin molecules in the crystallographic dimer suggest that the homogeneous post-translational modifications are an important factor in crystal formation.

Purification. Cell pellets were solubilized for 1 h at 4 °C with PBS buffer containing protease inhibitor tablets (complete protease inhibitor cocktail tablets, Roche) and 1.25% DM (β-decylmaltoside). Nuclei and other unsolubilized material were removed by centrifugation and the supernatant incubated with 1D4 antibody coupled to CNBr-activated sepharose (Amersham Biosciences). After 3–4 h the matrix was washed with PBS, pH 7, and 0.125% DM. Ground-state rhodopsin was reconstituted by adding 11-*cis* retinal (50 µM) to the matrix and overnight incubation at 4 °C. All steps involving retinal were performed under dim red light.

The matrix was washed with PBS, pH 7.0, and 0.125% DM, followed by 10 mM Hepes, pH 7.0, and 1% OG (β-octylglycoside). The purified protein was eluted in the same buffer supplemented with an elution peptide resembling the C terminus of rhodopsin (TETSQVAPA, 80 µM). The eluate was reduced to 0.5 ml using a 50-kD cut-off concentrator before gel filtration on a Superdex200 column. The gel filtration step was used to exchange the buffer for 10 mM Mes, pH 5.0, 100 mM NaCl, and 1% OG, which leads to protonation of the Schiff base in the Glu 113 Gln counterion mutant and shifts the absorption maximum from 382 to 498 nm (refs 33–35).

Light activation and crystallization. Reconstituted Glu 113 Gln was concentrated to 5–7.5 mg ml^{−1} (Vivaspin, 50-kD cut-off concentrator) and mixed with dried brain lipid extract (Avanti Polar Lipids, 1 w/w, 10 µM all-*trans* retinal). The sample was briefly sonicated and incubated for 30 min in presence of a tenfold molar excess of peptide resembling the last 11 amino acids of the Gα subunit of the G protein carrying the mutation Lys 341 Leu (ILENLKDCGLF, Advanced Biomedical). The receptor was activated for 5 min using a >515-nm long-pass filter that prevented light exposure of free retinal and metarhodopsin II.

The sample was mixed 1:1 with 3.0–3.4 M ammonium sulphate, 100 mM sodium acetate, pH 4.5, and crystallized by sitting-drop vapour diffusion in the dark. Crystals were collected under dim red light and soaked in crystallization buffer containing 10% trehalose before freezing in liquid nitrogen.

Data collection and structure determination. The diffraction quality of the crystals was assessed at several synchrotron X-ray sources (the Swiss Light Source, the European Synchrotron Radiation Facility and the Diamond Light Source). The best data with a maximum resolution of 2.9 Å was collected at the microfocus beamline ID23eh-2 at the European Synchrotron Radiation Facility by the helical data collection method. Data were integrated using XDS³⁶ and brought onto a common scale using SCALA from the CCP4 program suite³⁷. Data set statistics are given in Supplementary Table 2.

Phases were obtained by molecular replacement using the program PHASER³⁸ and the polypeptide of the opsin structures (PDB IDs, 3CAP³⁹ and 3DQB⁴⁰) or the ground-state rhodopsin structure (PDB ID, 1GZM⁴¹) as search model. The resulting solution was refined using iterative cycles of model building in COOT⁴² and refinement (rigid body, energy minimization, simulated annealing, individual B-factor refinement) with the PHENIX program suite⁴³. Water molecules were added to the model on the basis of five criteria: the presence of clear difference peaks in a 2.5σ-contoured $F_o - F_c$ electron density map calculated after simulated annealing refinement in which water molecules have been omitted (Fig. 3); density above 1σ in $2F_o - F_c$ maps refined with the inclusion of water molecules; two or more hydrogen bonds to the protein or other water molecules; a B-factor cut-off within 20% of the average; and interpretable density in maps calculated with both data sets.

Geometric restraints for lipids and heteroatoms were prepared using the PRODRG2 server (<http://davapcl.bioch.dundee.ac.uk/prodrg/>) and HIC-Up database (<http://xray.bmc.uu.se/hicup/>). Coordinates for all-*trans* and 9-*cis* retinal were obtained from the Cambridge Small Molecule Database. Retinal geometry restraints used in the refinement were prepared by carefully adjusting torsion angles and planarity restraints in the retinal parameter file distributed as part of the CCP4 program suite.

- Reeves, P. J., Kim, J. M. & Khorana, H. G. Structure and function in rhodopsin: a tetracycline-inducible system in stable mammalian cell lines for high-level expression of opsin mutants. *Proc. Natl Acad. Sci. USA* **99**, 13413 (2002).
- Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous *N*-glycosylation by a tetracycline-inducible *N*-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl Acad. Sci. USA* **99**, 13419 (2002).
- Sakmar, T. P., Franke, R. R. & Khorana, H. G. Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. *Proc. Natl Acad. Sci. USA* **86**, 8309 (1989).
- Zhukovsky, E. A. & Oprian, D. D. Effect of carboxylic acid side chains on the absorption maximum of visual pigments. *Science* **246**, 928 (1989).
- Standfuss, J., Zaitseva, E., Mahalingam, M. & Vogel, R. Structural impact of the E113Q counterion mutation on the activation and deactivation pathways of the G protein-coupled receptor rhodopsin. *J. Mol. Biol.* **380**, 145 (2008).
- Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795 (1993).
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* **61**, 458–464 (2005).
- Park, J. H. *et al.* Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454**, 183 (2008).
- Scheerer, P. *et al.* Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **455**, 497 (2008).
- Li, J. *et al.* Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **343**, 1409 (2004).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).

Amygdala circuitry mediating reversible and bidirectional control of anxiety

Kay M. Tye^{1*}, Rohit Prakash^{1,2*}, Sung-Yon Kim^{1,2*}, Lief E. Fenno^{1,2*}, Logan Groseknick^{1,2}, Hosniya Zarabi¹, Kimberly R. Thompson¹, Viviana Gradinaru^{1,2}, Charu Ramakrishnan¹ & Karl Deisseroth^{1,2,3,4,5}

Anxiety—a sustained state of heightened apprehension in the absence of immediate threat—becomes severely debilitating in disease states¹. Anxiety disorders represent the most common of psychiatric diseases (28% lifetime prevalence)² and contribute to the aetiology of major depression and substance abuse^{3,4}. Although it has been proposed that the amygdala, a brain region important for emotional processing^{5–8}, has a role in anxiety^{9–13}, the neural mechanisms that control anxiety remain unclear. Here we explore the neural circuits underlying anxiety-related behaviours by using optogenetics with two-photon microscopy, anxiety assays in freely moving mice, and electrophysiology. With the capability of optogenetics^{14–16} to control not only cell types but also specific connections between cells, we observed that temporally precise optogenetic stimulation of basolateral amygdala (BLA) terminals in the central nucleus of the amygdala (CeA)—achieved by viral transduction of the BLA with a codon-optimized channelrhodopsin followed by restricted illumination in the downstream CeA—exerted an acute, reversible anxiolytic effect. Conversely, selective optogenetic inhibition of the same projection with a third-generation halorhodopsin¹⁵ (eNpHR3.0) increased anxiety-related behaviours. Importantly, these effects were not observed with direct optogenetic control of BLA somata, possibly owing to recruitment of antagonistic downstream structures. Together, these results implicate specific BLA–CeA projections as critical circuit elements for acute anxiety control in the mammalian brain, and demonstrate the importance of optogenetically targeting defined projections, beyond simply targeting cell types, in the study of circuit function relevant to neuropsychiatric disease.

Despite the high prevalence^{1,2} of anxiety disorders, the underlying neural circuitry is incompletely understood. Available treatments are inconsistently effective or, in the case of benzodiazepines, addictive and linked to significant side effects including cognitive impairment and respiratory suppression¹⁷, pointing to the need for a deeper understanding of anxiety control mechanisms in the mammalian brain.

Although amygdala microcircuitry for conditioned fear has been optogenetically dissected^{18,19}, the causal underpinnings of unconditioned anxiety¹¹ have not yet been investigated with cellular precision. Pointing to the need for precise optogenetic exploration, the amygdala is composed of functionally and morphologically heterogeneous subnuclei with complex interconnectivity. The BLA is primarily glutamatergic (~90%)^{20,21} whereas the CeA, which encompasses the centrolateral (CeL) and centromedial (CeM) nuclei, consists of ~95% GABAergic medium spiny neurons²². The primary output region of the amygdala is the CeM^{23,24}, which (when chemically or electrically excited) mediates autonomic and behavioural responses associated with fear and anxiety via projections to the brainstem²⁵. Because patients with generalized anxiety disorder may have abnormal activity arising from the BLA and CeM¹¹, and as BLA neurons excite GABAergic CeL neurons²⁶ that

provide feed-forward inhibition onto CeM ‘output’ neurons^{6,18}, we considered that the BLA–CeL–CeM circuitry could be causally involved in anxiety. However, BLA pyramidal neurons as a whole could have varied and antagonistic roles in diverse projections throughout the brain, with targets including the bed nucleus of the stria terminalis (BNST), nucleus accumbens, hippocampus and cortex²⁶.

We therefore developed a method to selectively control BLA terminals in the CeA (Supplementary Methods). BLA glutamatergic projection neurons were transduced with an adeno-associated virus serotype 5 (AAV5) carrying codon-optimized channelrhodopsin (ChR2)–eYFP under control of the CaMKII α promoter followed by unilateral implantation of a bevelled guide cannula to allow preferential illumination of the non-transduced CeL (Supplementary Figs 1, 2). *In vivo* electrophysiological recordings were used to determine illumination parameters for selective control of those BLA terminals in the CeA without nonspecific control of all BLA somata (Supplementary Fig. 3).

To investigate the functional role of the BLA–CeA pathway in anxiety, we probed freely moving mice under projection-specific optogenetic control in two well-validated²⁷ anxiety assays: the elevated-plus maze (EPM) and the open-field test (OFT) (Fig. 1a–e). Mice display anxiety-related behaviours in open spaces; therefore, increased time spent in the open arms of the EPM or in the centre of the OFT is interpreted as reduced anxiety²⁷. To test whether anxiety-related behaviours could be related to activation of the BLA–CeA projection, and not all BLA somata as a whole, we compared mice receiving projection-specific photostimulation (ChR2:BLA–CeA; Fig. 1a) to a group with identical illumination parameters transduced with a control virus (eYFP:BLA–CeA), and to another control group expressing ChR2 in the BLA receiving direct illumination of the BLA (ChR2:BLA(somata)). Photostimulation of BLA terminals in the CeA (ChR2:BLA–CeA) increased open-arm time ($F_{1,42} = 69.09$, $P < 0.00001$; Fig. 1b, c) and the probability of open-arm entry from the maze centre ($F_{1,42} = 24.69$, $P < 0.00001$; Fig. 1c, inset, and Supplementary Movie) on the EPM, as well as increasing centre time in the OFT ($F_{1,105} = 24.46$, $P < 0.00001$; Fig. 1d, e), reflecting anxiety reduction without impaired locomotion (Supplementary Fig. 4). In contrast, the ChR2:BLA(somata) group showed reduced open-arm time ($F_{1,42} = 6.20$, $P = 0.02$; Fig. 1b, c) and probability of open-arm entry ($F_{1,42} = 5.15$, $P = 0.03$) during photostimulation relative to eYFP:BLA–CeA controls, reflecting a distinct anxiogenic effect. Thus, selective illumination of BLA projections to the CeA, but not of BLA somata nonspecifically, produced an acute, rapidly reversible anxiolytic effect.

Next we investigated the physiological basis of this light-induced anxiolysis. We considered that preferential photostimulation of BLA terminals in the CeL could activate CeL neurons and exert feed-forward inhibition onto brainstem-projecting CeM output neurons¹⁸ to implement anxiolysis. To test this, we undertook *in vivo* experiments, with light-delivery protocols matched to those delivered in the behavioural experiments, using activity-dependent immediate early

¹Department of Bioengineering, Stanford University, Stanford, California 94305, USA. ²Neurosciences Program, Stanford University, Stanford, California 94305, USA. ³Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, California 94305, USA. ⁴Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, USA. ⁵CNC Program, Stanford University, Stanford, California 94305, USA.

*These authors contributed equally to this work.

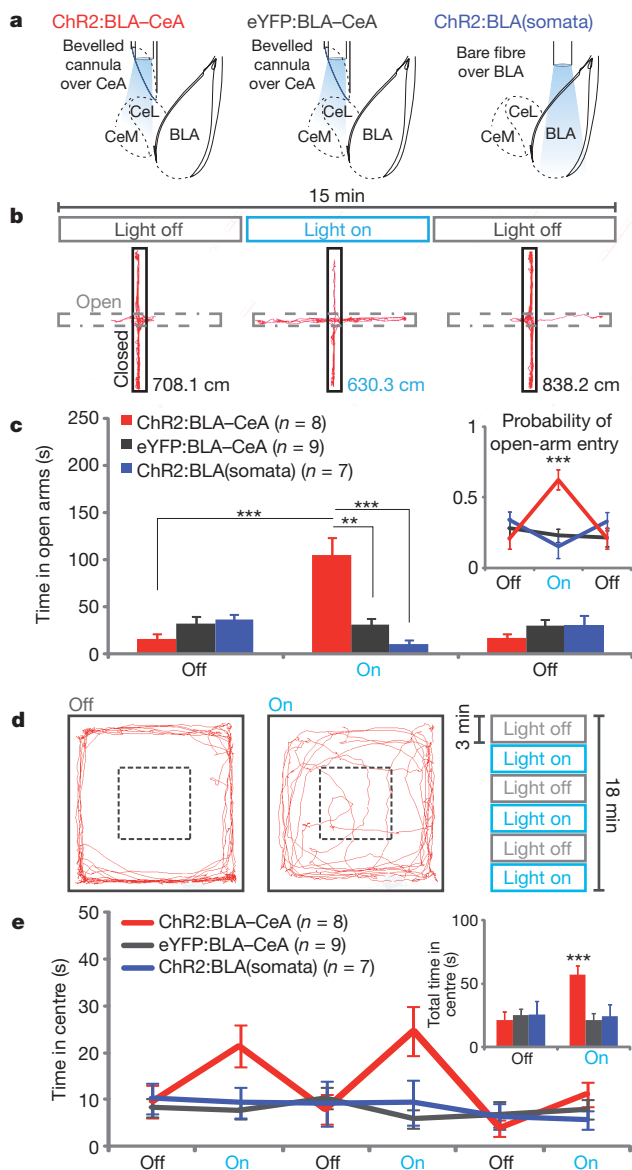


Figure 1 | Projection-specific excitation of BLA terminals in the CeA induces acute reversible anxiolysis. **a**, Mice were housed in a high-stress environment before behavioural manipulations, and received 5-ms light pulses at 20 Hz for all light-on conditions. **b**, **c**, ChR2:BLA-CeA mice ($n = 8$) received selective illumination of BLA terminals in the CeA during the light-on epoch on the EPM; see ChR2:BLA-CeA representative path (**b**), which induced an increase in open-arm time on photostimulation relative to eYFP:BLA-CeA ($n = 9$) and ChR2:BLA(somata) ($n = 7$) controls (**c**), and an increase in the probability of open-arm entry (see inset). **d**, **e**, ChR2:BLA-CeA mice also showed increased centre time on the OFT, as seen in a representative path (**d**), during light-on epochs relative to light-off epochs and eYFP:BLA-CeA and ChR2:BLA(somata) controls (**e**).

gene (*c-fos*) expression to track neuronal activation. We quantified the proportion of neurons in the BLA, CeL and CeM (Supplementary Fig. 5) within each group that expressed eYFP or showed *c-fos* immunoreactivity. Opsin expression was specific to BLA glutamatergic neurons, and was not observed in intercalated cells (Supplementary Fig. 6). No group differences were detected in the proportion of eYFP-positive cells within each region (Supplementary Fig. 5). We found a significantly higher proportion of *c-fos*-positive BLA cells in the ChR2:BLA(somata) group ($F_{2,9} = 10.12$, $P < 0.01$), relative to the ChR2:BLA-CeA or eYFP:BLA-CeA groups ($P < 0.01$ and $P < 0.05$, respectively), but no detectable difference between the ChR2:BLA-CeA and eYFP:BLA-CeA groups, indicating that the bevelled cannula shielding effectively

prevented BLA somata photostimulation. A higher proportion of CeL neurons expressed *c-fos* in the ChR2:BLA-CeA group relative to the eYFP:BLA-CeA group ($F_{2,9} = 4.54$, $P = 0.04$), but not the ChR2:BLA(somata) group (Supplementary Fig. 5). Thus, the *in vivo* illumination of BLA-CeA projections that triggered acute anxiolysis was found to excite CeL neurons without activating BLA somata.

To test the hypothesis that preferential illumination of BLA-CeL terminals induced feed-forward inhibition of CeM output neurons, we combined whole-cell patch-clamp electrophysiology with two-photon imaging to visualize the microcircuit while simultaneously probing the functional relationships among these cells during projection-specific optogenetic control (Fig. 2a–f). BLA neurons showed high-fidelity spiking to direct illumination (Fig. 2b). Illumination of BLA terminals in CeL elicited excitatory responses with stable spiking fidelity, and responding cells included both weakly and strongly excited CeL cells ($n = 16$; Fig. 2c).

To test whether illumination of BLA-CeL synapses could block CeM spiking via feed-forward inhibition from CeL neurons, we recorded from CeM neurons while selectively illuminating BLA-CeL synapses (Fig. 2d). Indeed, we observed potent spiking inhibition in CeM cells on illumination of BLA-CeL terminals (Fig. 2d; $F_{2,11} = 15.35$, $P = 0.004$). Figure 2e shows CeM responses recorded during illumination of ChR2-expressing BLA neurons in the CeM; importantly, the very same CeM neurons ($n = 7$) showed net excitation with broad illumination of BLA inputs to the CeM (Fig. 2e), but showed net inhibition on selective illumination of BLA inputs to the CeL (Fig. 2f). These data from a structurally and functionally identified microcircuit²⁵ illustrate that the balance of direct and indirect inputs from the BLA to the CeM can modulate CeM activity. We then directly tested whether overlapping or distinct populations of BLA neurons projected to the CeL and CeM in the mouse (Supplementary Fig. 7) by two-photon imaging in 350- μm -thick coronal slices. Of the BLA neurons sampled ($n = 18$; Supplementary Fig. 7), 44% projected to the CeL alone and 17% projected to the CeM alone, with only one cell observed to project to both the CeL and the CeM.

Because *in vivo* *c-fos* assays had indicated that illumination of BLA terminals in the CeL sufficed to drive excitation of postsynaptic CeL neurons but not to recruit efficiently BLA neurons as a whole, we next investigated the properties of optogenetically driven terminal stimulation in this microcircuit using whole-cell recordings. We first recorded from BLA pyramidal neurons expressing ChR2 and moved a restricted light spot ($\sim 125\ \mu\text{m}$ in diameter) in 100- μm steps from the cell soma, both in a direction over the visually identified axon and in a direction where there was no axon collateral, illustrating the spatial properties of light scattering in this circuit (Fig. 3a, b); spiking fidelity in the BLA neuron and evoked inward currents are summarized (Fig. 3a). Next, we found that typical photostimulation parameters drove reliable transmission when delivered to BLA-CeA synapses (assayed with recordings from postsynaptic CeL neurons; Fig. 3b and Supplementary Fig. 8); in contrast, when recording from the BLA somata instead, $\sim 300\ \mu\text{m}$ from the light spot, we did not observe reliable antidromically driven action potential firing (only $\sim 5\%$ reliability) despite use of the very same BLA-CeA synapse illumination conditions that elicited 100% reliable transmitter release from the illuminated terminals and the same cells known to spike robustly in response to somatic illumination (Fig. 3b). These results were consistent with the *c-fos* immunoreactivity and behavioural data (Fig. 1) and held even on bath application of GABA and glutamate receptor antagonists ($n = 7$) to eliminate local circuitry effects. The marked difference between effective synaptic transmission and antidromic spiking fidelity ($P = 0.0039$; Fig. 3b, inset) reveals that optogenetically driven vesicle release can occur in the absence of reliable antidromic drive, a potentially useful property that may relate to projection parameters such as axon calibre and myelination status (optogenetic stimuli will recruit thinner axons more efficiently than electrical stimuli), as well as experimental light intensity and spatial restriction properties.

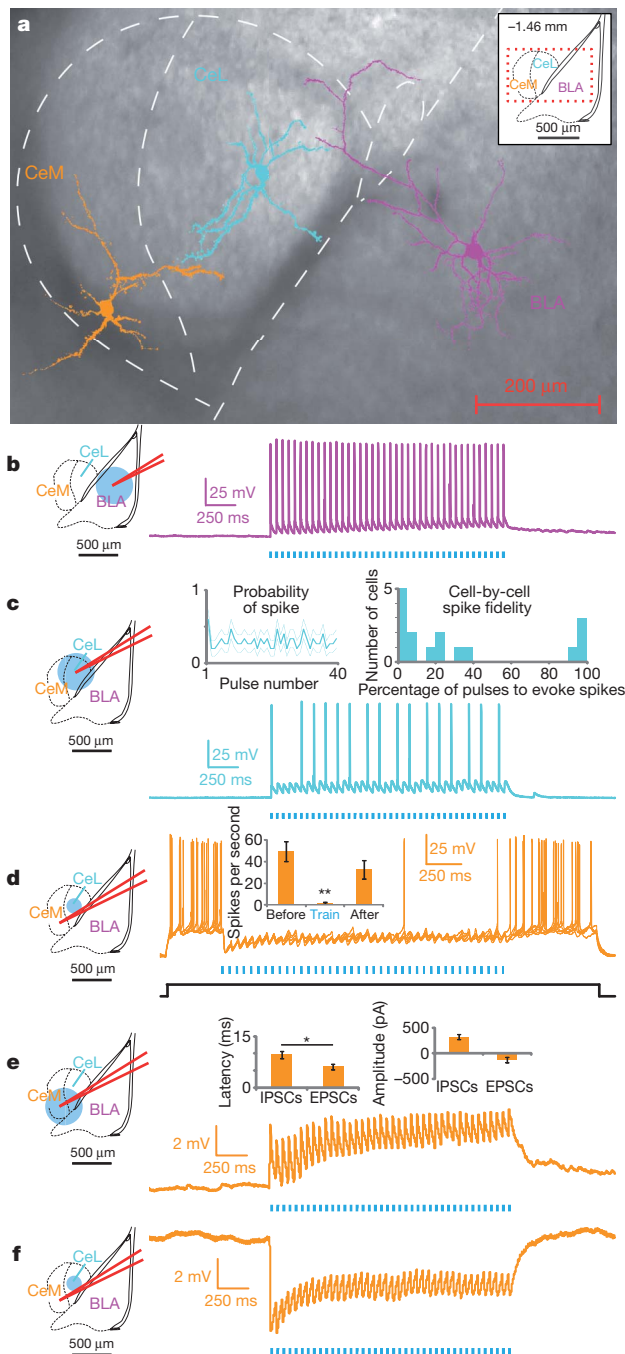


Figure 2 | Projection-specific excitation of BLA terminals in the CeA activates CeL neurons and elicits feed-forward inhibition of CeM neurons.

a, Two-photon images of representative BLA, CeL and CeM cells imaged from the same slice, overlaid on a brightfield image. **b–f**, Schematics of the recording and illumination sites for the associated representative current-clamp traces (membrane potential $V_m = \sim -70$ mV). **b**, Representative BLA pyramidal neuron trace expressing ChR2, all of which spiked for every pulse ($n = 4$). **c**, Representative trace from a CeL neuron in the terminal field of BLA projection neurons, showing both sub- and suprathreshold excitatory responses on photostimulation ($n = 16$). Inset left, population summary of mean probability of spiking for each pulse in a 40-pulse train at 20 Hz, dotted lines indicate s.e.m. Inset right, frequency histogram showing individual cell spiking fidelity; y-axis is the number of cells per each 5% bin. **d**, Six sweeps from a CeM neuron spiking in response to a current step (~ 60 pA; indicated in black) and inhibition of spiking on 20 Hz illumination of BLA terminals in the CeL. Inset, spike frequency was significantly reduced during light stimulation of CeL neurons ($n = 4$; spikes per second before (49 ± 9.0), during (1.5 ± 0.87) and after (33 ± 8.4) illumination; mean \pm s.e.m.). **e, f**, On broad illumination of the CeM, voltage-clamp summaries show that the latency of excitatory postsynaptic currents (EPSCs) is significantly shorter than the latency of inhibitory postsynaptic currents (IPSCs), whereas there was a non-significant difference in the amplitude of EPSCs and IPSCs ($n = 11$; $*P = 0.04$, see insets). The same CeM neurons ($n = 7$) showed either net excitation when receiving illumination of the CeM (**e**) or net inhibition on selective illumination of the CeL (**f**).

drug treatment did not impair locomotor activity (Supplementary Fig. 9), and in acute slices time-locked light-evoked excitatory responses were abolished on bath application of 2,3-dihydroxy-6-nitro-7-sulfamoyl-benzo[f]quinoxaline-2,3-dione (NBQX) and (2R)-amino-5-phosphonovaleric acid; (2R)-amino-5-phosphonopentanoate (AP5) (Supplementary Fig. 10). These data demonstrate that the light-induced anxiolytic effects were caused by the activation of BLA–CeA synapses, and not attributable to BLA projections to distal targets passing through the CeA.

Finally, to test whether basal anxiety-reducing processes could be blocked by selectively inhibiting the BLA–CeA pathway, we bilaterally transduced either eNpHR3.0—which hyperpolarizes neuronal membranes on illumination with amber light¹⁵—or eYFP alone, under the CaMKII α promoter in the BLA, and implanted bilateral bevelled guide cannulae to allow selective illumination of BLA terminals in the CeA (Supplementary Fig. 11). eNpHR3.0 expression was restricted to glutamatergic neurons in the BLA, and the eNpHR3.0:BLA–CeA group showed raised levels of *c-fos* expression relative to the eYFP:BLA–CeA and eNpHR3.0:BLA(somata) groups in the CeM ($P < 0.05$; Supplementary Fig. 12), consistent with the hypothesis that inhibition of BLA–CeL synapses suppresses feed-forward inhibition from CeL neurons to CeM neurons, thereby increasing CeM excitability and the downstream processes leading to increased anxiety phenotypes. Selective illumination of eNpHR3.0-expressing axon terminals reduced the probability of both spontaneously occurring (frequency: $F_{1,8} = 32.99$, $P = 0.00024$; amplitude: $F_{1,8} = 21.96$, $P = 0.001$; Supplementary Fig. 13) and electrically evoked ($F_{1,10} = 10.79$, $P = 0.006$; Fig. 4a) vesicle release, without preventing spiking at the soma (Supplementary Fig. 14). BLA somata inhibition did not induce an anxiogenic response, perhaps owing to the simultaneous decrease in direct BLA–CeM excitatory input. We also found that the eNpHR3.0:BLA–CeA group showed reduced open-arm time and probability of open-arm entry on the EPM ($F_{1,40} = 21.08$, $P < 0.00001$; and $F_{1,40} = 19.93$, $P < 0.00001$, respectively; Fig. 4b, c) and centre time in the OFT ($F_{1,100} = 18.919$, $P < 0.00001$; Fig. 4d, e) during photostimulation when compared to the eYFP and BLA(somata) groups, without altering locomotor activity (Supplementary Fig. 15). These data demonstrate that preferential inhibition of BLA–CeL synapses acutely increases anxiety-like behaviours.

Here, we have identified the BLA–CeL pathway as a neural substrate for real-time bidirectional modulation of the unconditioned expression of anxiety. The observation that selective illumination of specific BLA terminals produces distinct, and even opposite, behavioural responses

To confirm further that the anxiolytic effect was due to the selective activation of BLA–CeL projections alone, and not BLA axons passing through the CeA or back-propagation of action potentials to BLA cell bodies that would then innervate all BLA projection target regions, we tested whether local glutamate receptor antagonism would attenuate light-induced anxiolytic effects. This distinction is critical, as there have been previous reports that CeA lesions that alter anxiety are confounded by the ablation of BLA projections to the BNST passing through the CeA²⁸. In a separate group of mice, we selectively illuminated BLA–CeA terminals as before ($n = 8$; Supplementary Fig. 1), but infused glutamate antagonists or saline via the fibre-optic guide cannula before testing on the EPM and OFT. Confirming a local synaptic mechanism rather than control of fibres of passage, intra-CeA glutamate receptor antagonism abolished light-induced reductions in anxiety as measured by open-arm time ($F_{1,35} = 8.61$, $P = 0.008$) and probability of open-arm entry on the EPM ($F_{1,35} = 5.92$, $P = 0.02$), and centre time during the OFT ($F_{1,77} = 13.99$, $P = 0.0006$; Fig. 3c, d). Importantly,

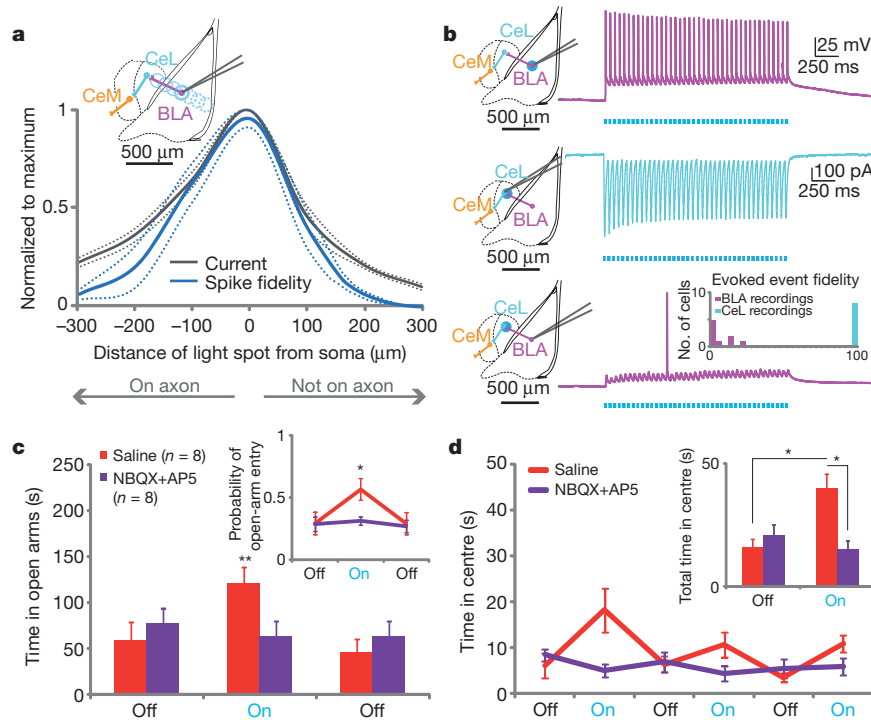


Figure 3 | Light-induced anxiolytic effects are attributable to activation of BLA–CeA synapses. **a, b**, Schematic of the recording site and illumination positions as whole-cell recordings were performed at each illumination location in 100- μ m increments away from the cell soma both over a visualized axon and in a direction that was not over an axon (inset). Normalized summary of spike fidelity and depolarizing current (a) to a 20-Hz train delivered at various distances from the soma. **b**, Representative traces on ~125- μ m-diameter illumination at various locations within each slice ($n = 7$). Illumination of BLA somata elicits high-fidelity spiking (top). Illumination of BLA terminals in CeL elicits strong excitatory responses shown in voltage-clamp in the postsynaptic

CeL neuron (middle), but does not elicit reliable antidromic spiking in the BLA neuron itself (bottom), summarized in a frequency histogram (inset, 120 pulses per cell). **c, d**, A separate group of ChR2:BLA–CeA mice ($n = 8$) performed the EPM and OFT twice, one session preceded with intra-CeA infusions of saline (red) and the other session with glutamate receptor antagonists NBQX and AP5 (purple), counter-balanced for order. Glutamate receptor blockade in the CeA attenuated light-induced increases in both open-arm time (c) and probability of open-arm entry (inset) on the EPM and centre time on the OFT (d, inset shows pooled summary), without altering baseline performance.

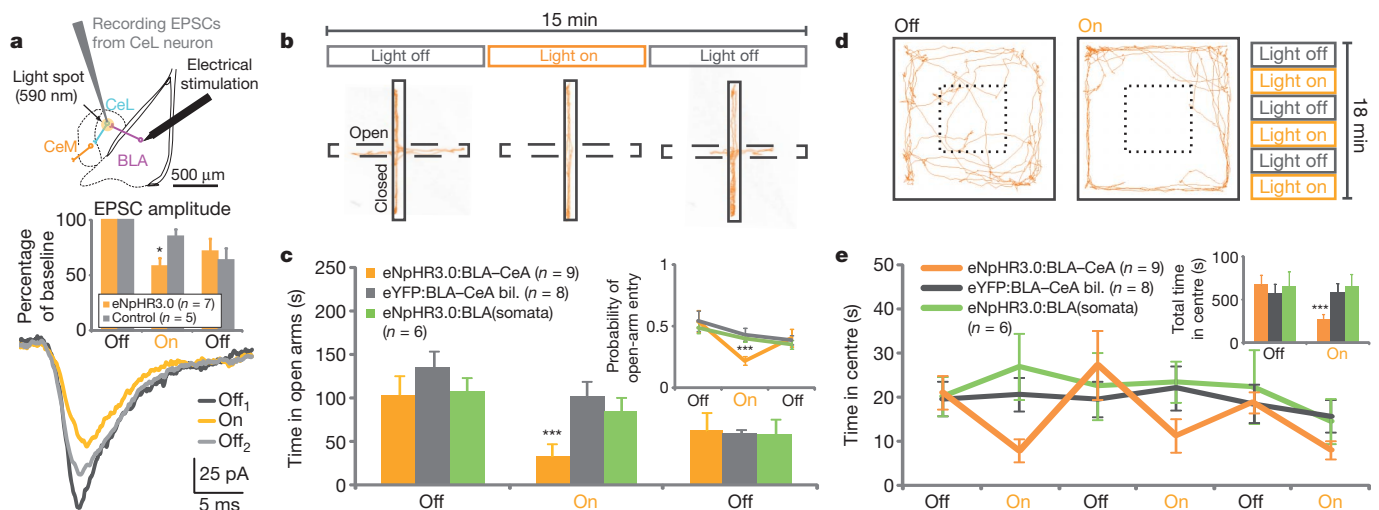


Figure 4 | Selective inhibition of BLA terminals in the CeA induces an acute and reversible increase in anxiety. **a–e**, Mice were group-housed in a low-stress environment and received bilateral constant 594-nm light during light-on epochs. **a**, Selective illumination of eNpHR3.0-expressing BLA terminals suppresses vesicle release evoked by electrical stimulation in the BLA. Schematic indicates the locations of the stimulating electrode, the recording electrode and the ~125- μ m diameter light spot. Representative CeL EPSCs before (Off₁), during (On) and after (Off₂) selective illumination of eNpHR3.0-expressing BLA terminals. Normalized EPSC amplitude summary data from sections containing BLA neurons expressing eNpHR3.0 ($n = 7$) and non-

transduced controls ($n = 5$) show that selectively illuminating BLA–CeL terminals reduces ($*P = 0.006$) electrically evoked EPSC amplitude in postsynaptic CeL neurons relative to non-transduced control slice preparations (inset). **b, c**, Representative eNpHR3.0:BLA–CeA path (b) indicates reduced open-arm time (c) and probability of open-arm entry (inset) during illumination, relative to controls. bil., bilateral. **d, e**, Representative eNpHR3.0:BLA–CeA path (d) reflects reduced centre time on the OFT (e) for the eNpHR3.0:BLA–CeA group during light-on, but not light-off, epochs as compared to controls (inset shows pooled data).

from illumination of all glutamatergic BLA somata nonspecifically, points to the essential value of optogenetic control in causally dissecting intact neural circuitry, and indicates that multiple subpopulations or projections of BLA neurons can act in opposition (for example, direct excitation of CeM along with feed-forward inhibition of CeM). Neural circuitry arranged in this way provides many opportunities for modulation of expression of anxiety phenotypes; for example, this microcircuit is well-positioned to be influenced by top-down cortical control from regions important for processing fear and anxiety, including the prelimbic, infralimbic, anterior cingulate and insular cortices that provide robust innervation to the BLA and CeL.

These data are consistent with reports implicating CeA involvement in anxiety^{9,11,12}, but it is important to note that our findings do not exclude downstream or parallel circuits including the BNST²⁸, the insular and prefrontal cortices²⁹, and the septal-hippocampal circuit³⁰; for example, stress induces CeL release of corticotropin releasing hormone (CRH) in the BNST²⁸. In the course of providing insight into native anxiogenic and anxiolytic processes, these findings demonstrate that anxiety is continuously regulated by balanced antagonistic pathways within the amygdala, and illustrate the importance of resolving specific projections in the study of neural circuit function relevant to psychiatric disease.

METHODS SUMMARY

Virus-mediated opsin gene expression. The pAAV-CaMKII α -hChR2(H134R)-eYFP, pAAV-CaMKII α -eYFP and pAAV-CaMKII α -eNpHR3.0-eYFP plasmids were designed and constructed by standard methods and packaged as AAV5. Virus (0.5 μ l) was injected into the BLA. Maps and clones are available at <http://www.optogenetics.org>.

In vivo projection-specific targeting. To investigate the role of the BLA–CeL pathway in modulating anxiety, we performed viral transduction and surgical implantation of bevelled guide cannulae to allow selective illumination of BLA fibres in the CeA under stereotaxic guidance. Behavioural, electrophysiological and imaging data were collected 4–6 weeks after surgery.

Two-photon imaging and functional mapping using ex vivo electrophysiology. Acute slices were collected for two-photon imaging and ex vivo electrophysiological recordings. While light-stimulation parameters used *in vivo* were delivered via fibre optics, and light in ex vivo experiments was delivered onto coronal sections, we matched light power density at our target region $\sim 6 \text{ mW mm}^{-2}$. Whole-cell recordings were made from BLA pyramidal neurons simultaneously during two-photon visualization of neuronal processes with Alexa Fluor dye. We visually tracked axonal projections from BLA neurons to the CeL nucleus. We recorded from CeL neurons on illumination with an aperture-restricted light spot ($\sim 125 \mu\text{m}$ diameter), mimicking the preferential illumination of BLA terminals, but not BLA somata, delivered *in vivo*. Two-photon imaging allowed axonal tracking to the CeM, where whole-cell recordings were collected from CeM neurons in the terminal field of CeL axons, with aperture-restricted illumination over the CeL to allow selective illumination of BLA terminals in the CeL while recording from the CeM neuron.

Opsin expression validation and immunohistochemistry. To validate specificity, sensitivity and spatial distribution of opsin expression as well as neuronal activity, brain slices were prepared for optical microscopy and immunohistochemistry. Coronal sections were stained for 4',6-diamidino-2-phenylindole (DAPI) and immunoreactivity for *c-fos*. Quantitative analyses of confocal images were performed with both staining and analysis blind to experimental condition.

Received 7 November 2010; accepted 14 January 2011.

Published online 9 March 2011.

1. Lieb, R. Anxiety disorders: clinical presentation and epidemiology. *Handb. Exp. Pharmacol.* **169**, 405–432 (2005).
2. Kessler, R. C. *et al.* Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* **62**, 593–602 (2005).
3. Koob, G. F. Brain stress systems in the amygdala and addiction. *Brain Res.* **1293**, 61–75 (2009).
4. Ressler, K. J. & Mayberg, H. S. Targeting abnormal neural circuits in mood and anxiety disorders: from the laboratory to the clinic. *Nature Neurosci.* **10**, 1116–1124 (2007).
5. LeDoux, J. The emotional brain, fear, and the amygdala. *Cell. Mol. Neurobiol.* **23**, 727–738 (2003).
6. Pare, D., Quirk, G. J. & LeDoux, J. E. New vistas on amygdala networks in conditioned fear. *J. Neurophysiol.* **92**, 1–9 (2004).

7. Tye, K. M., Stuber, G. D., de Ridder, B., Bonci, A. & Janak, P. H. Rapid strengthening of thalamo-amygdala synapses mediates cue-reward learning. *Nature* **453**, 1253–1257 (2008).
8. Weiskrantz, L. Behavioral changes associated with ablation of the amygdaloid complex in monkeys. *J. Comp. Physiol. Psychol.* **49**, 381–391 (1956).
9. Kalin, N. H., Shelton, S. E. & Davidson, R. J. The role of the central nucleus of the amygdala in mediating fear and anxiety in the primate. *J. Neurosci.* **24**, 5506–5515 (2004).
10. Lesscher, H. M. *et al.* Amygdala protein kinase C epsilon regulates corticotropin-releasing factor and anxiety-like behavior. *Genes Brain Behav.* **7**, 323–333 (2008).
11. Etkin, A., Prater, K. E., Schatzberg, A. F., Menon, V. & Greicius, M. D. Disrupted amygdalar subregion functional connectivity and evidence of a compensatory network in generalized anxiety disorder. *Arch. Gen. Psychiatry* **66**, 1361–1372 (2009).
12. Lyons, A. M. & Thiele, T. E. Neuropeptide Y conjugated to saporin alters anxiety-like behavior when injected into the central nucleus of the amygdala or basomedial hypothalamus in BALB/cJ mice. *Peptides* **31**, 2193–2199 (2010).
13. Roozendaal, B., McEwen, B. S. & Chattarji, S. Stress, memory and the amygdala. *Nature Rev. Neurosci.* **10**, 423–433 (2009).
14. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neurosci.* **8**, 1263–1268 (2005).
15. Gradinaru, V. *et al.* Molecular and cellular approaches for diversifying and extending optogenetics. *Cell* **141**, 154–165 (2010).
16. Deisseroth, K. Optogenetics: controlling the brain with light. *Sci. Am.* **303**, 48–55 (2010).
17. Woods, J. H., Katz, J. L. & Winger, G. Benzodiazepines: use, abuse, and consequences. *Pharmacol. Rev.* **44**, 151–347 (1992).
18. Cioocchi, S. *et al.* Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature* **468**, 277–282 (2010).
19. Haubensak, W. *et al.* Genetic dissection of an amygdala microcircuit that gates conditioned fear. *Nature* **468**, 270–276 (2010).
20. Carlsen, J. Immunocytochemical localization of glutamate decarboxylase in the rat basolateral amygdaloid nucleus, with special reference to GABAergic innervation of amygdalostratial projection neurons. *J. Comp. Neurol.* **273**, 513–526 (1988).
21. Smith, Y. & Pare, D. Intra-amygdaloid projections of the lateral nucleus in the cat: PHA-L anterograde labeling combined with postembedding GABA and glutamate immunocytochemistry. *J. Comp. Neurol.* **342**, 232–248 (1994).
22. McDonald, A. J. Cytoarchitecture of the central amygdaloid nucleus of the rat. *J. Comp. Neurol.* **208**, 401–418 (1982).
23. Krettek, J. E. & Price, J. L. A description of the amygdaloid complex in the rat and cat with observations on intra-amygdaloid axonal connections. *J. Comp. Neurol.* **178**, 255–279 (1978).
24. Krettek, J. E. & Price, J. L. Amygdaloid projections to subcortical structures within the basal forebrain and brainstem in the rat and cat. *J. Comp. Neurol.* **178**, 225–253 (1978).
25. Davis, M. in *The Amygdala: A Functional Analysis* (ed. Aggleton, J. P.) 213–288 (Oxford Univ. Press, 2000).
26. Pitkanen, A. in *The Amygdala: A Functional Analysis* (ed. Aggleton, J. P.) 31–99 (Oxford Univ. Press, 2000).
27. Carola, V., D'Olimpio, F., Brunamonti, E., Mangia, F. & Renzi, P. Evaluation of the elevated plus-maze and open-field tests for the assessment of anxiety-related behaviour in inbred mice. *Behav. Brain Res.* **134**, 49–57 (2002).
28. Davis, M., Walker, D. L., Miles, L. & Grillon, C. Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology* **35**, 105–135 (2010).
29. Shin, L. M. & Liberzon, I. The neurocircuitry of fear, stress, and anxiety disorders. *Neuropsychopharmacology* **35**, 169–191 (2010).
30. Gray, J. A. & McNaughton, N. The neuropsychology of anxiety: reprise. *Nebr. Symp. Motiv.* **43**, 61–134 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank P. Janak, H. Fields, G. Stuber, E. Thomas, F. Zhang, I. Witten, V. Sohal, T. Davidson and M. Warden as well as J. Mattis, R. Durand, M. Mogri, J. Mirzabekov and E. Steinberg for discussions, and the entire K.D. laboratory for their support. All viruses were packaged at University of North Carolina (UNC) Vector Core. Supported by NIMH (1F32MH088010-01, K.M.T.), NARSAD (K.R.T.), Samsung Scholarship (S.-Y.K.), NSF IGERT Award 0801700 (L.G.) and the Defense Advanced Research Projects Agency Reorganization and Plasticity to Accelerate Injury Recovery (N66001-10-C-2010), the Alice Woo, Albert Yu, Snyder, and McKnight Foundations, as well as NIDA, NIMH and the NIH Pioneer Award (K.D.).

Author Contributions K.M.T., R.P., S.-Y.K., L.E.F. and K.D. contributed to study design and data interpretation. K.M.T., R.P., S.-Y.K. and L.E.F. contributed to data collection and K.M.T. coordinated data collection and analysis. K.M.T., S.-Y.K., H.Z. and K.R.T. contributed to immunohistochemical processing, fluorescence imaging and quantitative analyses. K.M.T. and L.G. performed the behavioural and ex vivo electrophysiology statistical analyses. V.G. and C.R. contributed to the design of eNpHR3.0. C.R. cloned all constructs and managed viral packaging processes. K.D. supervised all aspects of the work. All authors contributed to writing the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.D. (deissero@stanford.edu).

Collapse of long-range charge order tracked by time-resolved photoemission at high momenta

Timm Rohwer^{1*}, Stefan Hellmann^{1*}, Martin Wiesenmayer¹, Christian Sohrt¹, Ankatrin Stange¹, Bartosz Slomski¹, Adra Carr², Yanwei Liu^{3,4}, Luis Mijang Avila⁵, Matthias Kalläne¹, Stefan Mathias^{2,6}, Lutz Kipp¹, Kai Rossnagel¹ & Michael Bauer¹

Intense femtosecond (10^{-15} s) light pulses can be used to transform electronic, magnetic and structural order in condensed-matter systems on timescales of electronic and atomic motion^{1,2,3}. This technique is particularly useful in the study^{4,5} and in the control⁶ of materials whose physical properties are governed by the interactions between multiple degrees of freedom. Time- and angle-resolved photoemission spectroscopy is in this context a direct and comprehensive, energy- and momentum-selective probe of the ultrafast processes that couple to the electronic degrees of freedom^{7–10}. Previously, the capability of such studies to access electron momentum space away from zero momentum was, however, restricted owing to limitations of the available probing photon energy^{10,11}. Here, using femtosecond extreme-ultraviolet pulses delivered by a high-harmonic-generation source, we use time- and angle-resolved photoemission spectroscopy to measure the photoinduced vaporization of a charge-ordered state in the potential excitonic insulator 1T-TiSe₂ (refs 12, 13). By way of stroboscopic imaging of electronic band dispersions at large momentum, in the vicinity of the edge of the first Brillouin zone, we reveal that the collapse of atomic-scale periodic long-range order happens on a timescale as short as 20 femtoseconds. The surprisingly fast response of the system is assigned to screening by the transient generation of free charge carriers. Similar screening scenarios are likely to be relevant in other photoinduced solid-state transitions and may generally determine the response times. Moreover, as electron states with large momenta govern fundamental electronic properties in condensed matter systems¹⁴, we anticipate that the experimental advance represented by the present study will be useful to study the ultrafast dynamics and microscopic mechanisms of electronic phenomena in a wide range of materials.

The electronic properties of many condensed matter systems are determined by large-momentum electron states, often located near the edge of the first Brillouin zone (BZ), the unit cell of crystalline solids in electron momentum (k) space. A prominent current example is graphene, whose hallmark, the critical crossing point of its peculiar conical band dispersion, is located at the corner of the BZ¹⁵. The copper oxide based high-temperature superconductors represent another example, where the much debated competition between the pseudogap and superconductivity is particularly pronounced in the so-called antinodal region of the Fermi surface, near the BZ boundary¹⁶. A final example is the new class of iron-based superconductors, which are characterized by magnetic excitations that couple two sets of Fermi surfaces, one set centred on the corners of the BZ¹⁷. More generally, it is the coupling of high momentum electron states near the Fermi momenta k_F (the momenta of the highest occupied electron states) that contributes most to the linear response function of an electron liquid, sometimes even causing a divergence that leads to phase instabilities.

Typical electron momenta near the boundary of the first BZ are in the 1 \AA^{-1} regime. Conventional angle-resolved photoemission spectroscopy (ARPES) with photon energies exceeding roughly 10 eV is probably the most powerful tool to map band structure peculiarities and Fermi surfaces up to and beyond these critical points: this technique can also, at the same time, determine many-body effects embodied in the fine details of band dispersions and in the distribution of spectral weight. ARPES is particularly well suited for layered materials—as in the present study—because for quasi-two-dimensional systems, the measured electronic structure can be considered as being predominantly characteristic of the bulk, despite the surface sensitivity of the probe. The great allure of corresponding time-resolved ARPES experiments is the provision of direct dynamical information and the possibility of disentangling—via temporal discrimination—the various interactions between the relevant degrees of freedom that determine material properties in the quantum world. Here, we present femtosecond time-resolved ARPES experiments, in which transient changes in the whole occupied electronic structure between the centre and the edge of the BZ are probed to answer a simple fundamental question: how fast can long-range charge order in a solid melt?

The charge-ordered state we investigate is the conspicuous ($2 \times 2 \times 2$) charge-density wave (CDW) that occurs in the layered compound 1T-TiSe₂ below a temperature of 200 K (ref. 13). Figure 1 presents the thermal equilibrium view of the CDW transition, which affects both structural and electronic properties. On the transition, the atoms move to new equilibrium positions such that the real-space unit cell doubles its size in all three directions (Fig. 1a and c; arrows in Fig. 1c indicate the atomic displacements from the normal-phase positions); in momentum space, the dimensions of the BZ are correspondingly halved (Fig. 1b and d). The new k -space geometry suggests that the wave vector of the CDW is determined by an interaction between the Se 4p valence band maximum at $\bar{\Gamma}$, the centre of the BZ, and the elliptical pocket of Ti 3d states at the BZ edge at \bar{M} (Fig. 1b). These symmetry points are connected by the new reciprocal lattice vectors and become equivalent $\bar{\Gamma}$ points in the new phase, which allows direct Se 4p–Ti 3d interaction (Fig. 1d). This interaction is in fact remarkably strong and extremely well resolved by conventional ARPES^{12,18,19}. Figure 1e and f compares ARPES intensity maps recorded with synchrotron radiation ($h\nu = 119 \text{ eV}$) along the $\bar{\Gamma}$ – \bar{M} direction above and below the CDW transition temperature. At $\bar{\Gamma}$, the downward dispersing (hole-like) Se 4p bands dominate the photoemission signal in both maps and only a small shift of the valence band maximum is visible. At \bar{M} , where the high-temperature map shows the bottom of the upward dispersing (electron-like) Ti 3d band, the effects are more dramatic: the CDW leads to a strong selective transfer of spectral weight by folding the Se 4p band from $\bar{\Gamma}$ onto \bar{M} . It is this remarkably strong folded Se 4p intensity which we will use in the time-resolved experiments as a spectroscopic measure for CDW order in 1T-TiSe₂.

¹Institute of Experimental and Applied Physics, University of Kiel, 24118 Kiel, Germany. ²JILA, University of Colorado and NIST, Boulder, Colorado 80309-0440, USA. ³University of California, Berkeley, NSF ERC Extreme Ultraviolet Science and Technology, Berkeley, California 94720, USA. ⁴Center for X-Ray Optics, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ⁵Center for Nano and Molecular Science, University of Texas at Austin, Austin, Texas 78722, USA. ⁶Department of Physics and Research Center OPTIMAS, University of Kaiserslautern, 67663 Kaiserslautern, Germany.

*These authors contributed equally to this work.

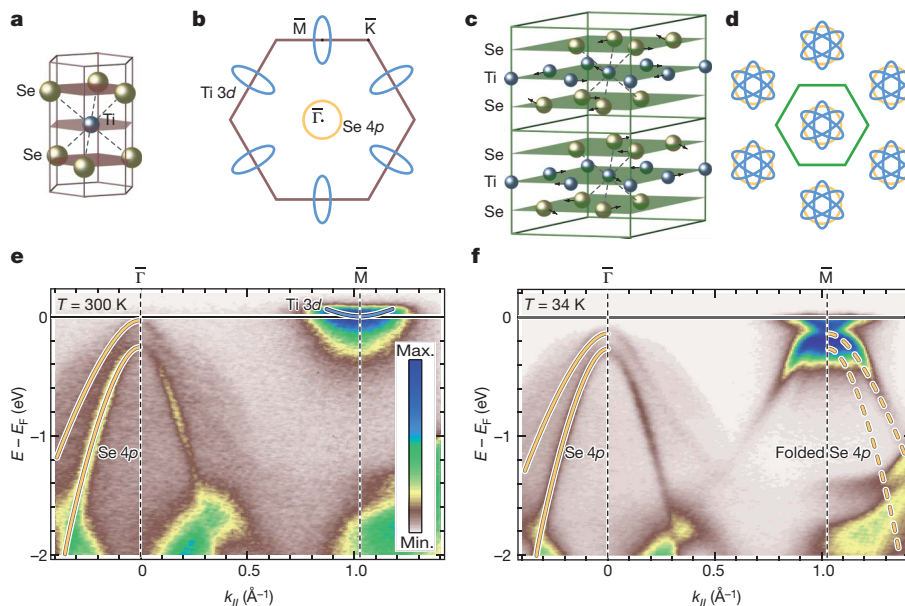


Figure 1 | CDW phase transition of 1T-TiSe₂. **a**, Real-space unit cell of the normal phase at room temperature. **b**, Common unit cell (first BZ, purple line) in momentum space of the room temperature phase (two-dimensional projection onto the surface plane). $\bar{\Gamma}$, \bar{M} and \bar{K} are high-symmetry points of the first BZ. Blue ellipses and yellow circle indicate the Fermi surface topology of Ti 3d and Se 4p bands, respectively. Planes show the atomic layers of the two-dimensional crystalline structure. **c**, Real-space unit cell of the CDW phase.

We now apply sub-10-fs extreme ultraviolet (XUV) pulses ($h\nu = 43$ eV, s-polarization) to monitor the transient response of the CDW phase to excitation with infrared laser pulses ($h\nu = 1.57$ eV) of 32 fs width. Details of the time-resolved ARPES experiment are described in Methods. Figure 2a shows ARPES intensity maps of 1T-TiSe₂ measured at $T = 125$ K with the femtosecond XUV light source²⁰ (Methods). Despite the poorer energy resolution, both sets of Se 4p bands are well resolved, the original one at $\bar{\Gamma}$ and the folded one at \bar{M} . Time-resolved experiments are performed at infrared pump fluences between 0.2 and 5 mJ cm⁻², corresponding to an excitation density range of 0.025 to 0.63 photons per Ti atom. Figure 2b–e shows four photoemission snapshots recorded at a pump fluence of 5 mJ cm⁻² with increasing temporal delay between the infrared pump and the XUV probe, up to a maximum of 3 ps (Supplementary Movie 1). The data have been corrected for a space charge shift of 200 meV induced by the electron background because of multi-photon photoemission by the infrared pump pulse. The time series is dominated by two prominent changes in the photoemission intensity maps. First, in instantaneous response to the infrared excitation, an electron-like band appears, crossing the Fermi energy E_F and extending (at sufficiently small temporal delays) from \bar{M} to $\bar{\Gamma}$. We observe here the transient generation of quasi-free charge carriers because of near-resonant Ti 3d – Se 4p excitation. Second, the downward-dispersing Se 4p band, folded onto the \bar{M} point owing to the interaction with the CDW superlattice, disappears or is at least considerably reduced in intensity. This suggests that long-range order in the electronic subsystem breaks down on an ultrafast timescale. In the following, we restrict our quantitative analysis to the short-time (sub-100-fs) dynamics of this process.

Figure 2f compares the temporal evolution of the integrated intensity of the folded Se 4p band—our spectroscopic measure for CDW order—for different pump fluences (see Supplementary Information section 3 for details of the data analysis). Both breakdown and (partial) recovery of the signal (inset of Fig. 2f) are strongly dependent on the pump fluence. The fluence dependence of the time constant characterising the signal breakdown, $\tau_{\text{Se } 4p}$, is shown in Fig. 3a: at the lowest fluences,

Arrows indicate the atomic displacements from the normal-phase positions; the dashed lines indicate the extension of the unit cell in the normal phase.

d, First BZ (green line) of the CDW phase. The folding of Se 4p and Ti 3d states is indicated. **e**, ARPES intensity map (electron binding energy versus momentum) of the room temperature phase. Photoelectron intensity is encoded in a false-colour scale. **f**, ARPES intensity map of the CDW phase.

the initial drop in the signal is retarded by about 80 fs with respect to the laser pulse excitation. As the fluence increases, the response becomes continuously faster, and at the highest fluences the transient minimum in the folded Se 4p band intensity appears well within the 32-fs-long infrared pump laser pulse with an ultimate response time of 20 fs. For comparison, the dynamics associated with the initial population of the Ti 3d band due to absorption happens within the width of the infrared pulse for the entire pump fluence regime. Notably, for the highest excitation fluence, the folded Se 4p band follows this population dynamics without delay (Supplementary Information section 4). The partial recovery of the folded Se 4p intensity is observed on timescales of several hundreds of femtoseconds. Two-temperature model calculations following reference 21 suggest that this recovery is mostly driven by thermalization of the electronic subsystem with the atomic lattice.

In previous studies, it has been shown that the fundamental timescales of photoinduced phase transitions are governed by bottlenecks associated with the characteristic response times of the relevant degrees of freedom, such as the oscillation period of neighbouring atoms²² or the hopping rate of localized electrons between neighbouring sites²³. The upper solid line in Fig. 3a marks for instance the expected short-time limit (75 fs) of the lattice response of 1T-TiSe₂ to a photoexcitation; this short-time limit is taken as one-quarter of the oscillation period of the high-frequency CDW amplitude mode²⁴. The vaporization of a long-range-ordered state within 20 fs is in this context exceptionally fast. The ultrafast timescale observed in the high-fluence regime points to a purely electronically driven process, whose response time, however, strongly depends on the excitation fluence. The absorption of the light pulse initially increases the free charge carrier density n (electrons and holes), as can be seen in the instantaneous population of the Ti 3d band. This transient free carrier population, which is directly governed by the excitation fluence, links the time constant $\tau_{\text{Se } 4p}$ to a material specific timescale: quantum kinetic calculations have shown that the characteristic build-up time for carrier screening in response to an ultrashort laser excitation is the plasma oscillation period, τ_{pl} (refs 25, 26), which scales with $1/\sqrt{n}$. As shown in Fig. 3a, $\tau_{\text{Se } 4p}$ closely follows such a $1/\sqrt{n}$ dependence (a quantitative estimate of the

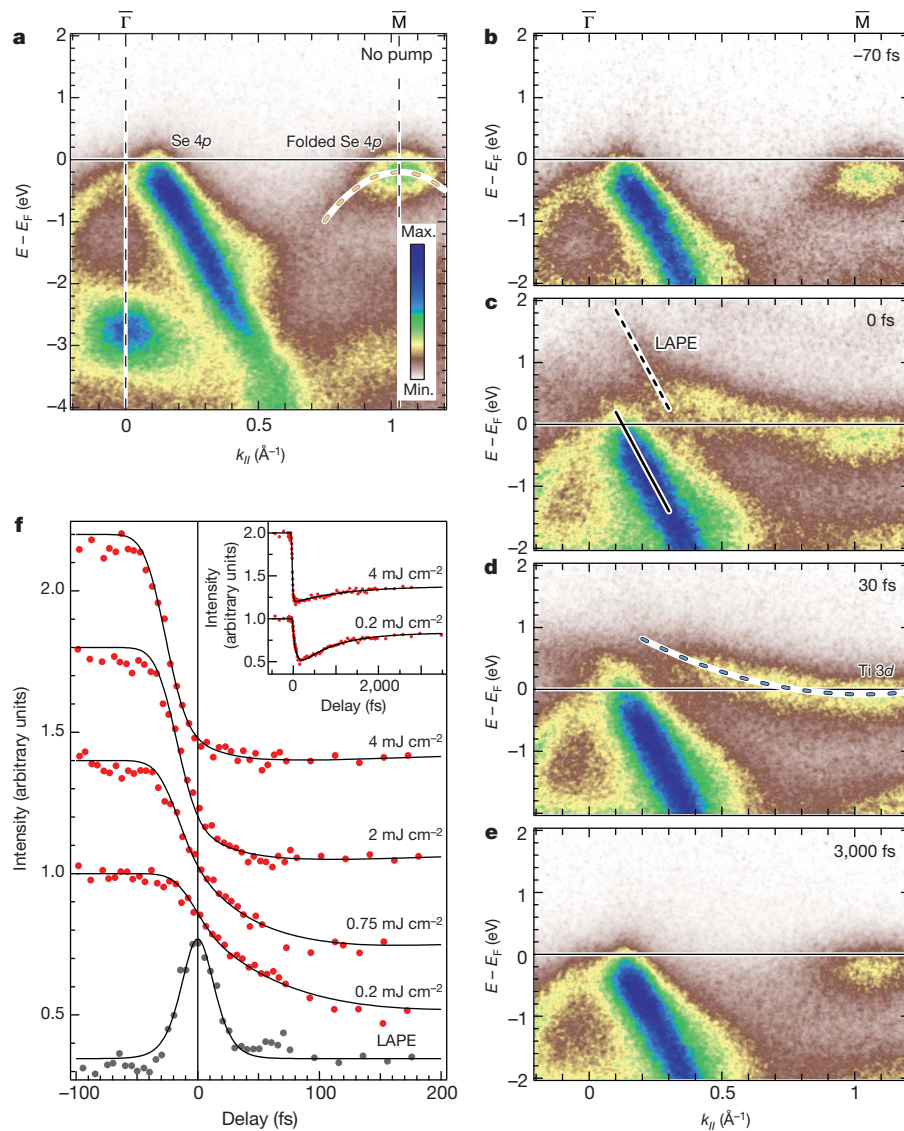


Figure 2 | Tracking the photoinduced transition by femtosecond time-resolved ARPES. **a**, ARPES intensity map of 1T-TiSe₂ recorded with high-harmonic XUV pulses in the CDW phase (temperature $T = 125$ K). **b–e**, Time-resolved ARPES snapshots at increasing pump (infrared)–probe (XUV) temporal delays (pump–pulse fluence, 5 mJ cm^{-2}). Energy distribution curves (EDCs) are provided in Supplementary Information 2. **f**, Photoemission transients of the folded Se 4p band for different infrared excitation fluences close to time zero. The infrared–XUV cross-correlation signal, which has been determined from the laser assisted photoemission³⁰ (LAPE) signal, is added (Methods). Inset, transients up to temporal delays of 3.5 ps.

photoinduced carrier density n and the corresponding plasma oscillation period is given in Supplementary Information section 5). It has in fact been shown that the build-up of screening by photo-injected carriers is relevant for the dynamics of ultrafast processes on the sub-100-fs timescale²⁷.

Suppression of screening is important in correlation-induced phase transitions. Therefore, it is not surprising that the reverse—the enhancement of screening by photo-injection of free carriers—can destroy a correlation-induced phase. The CDW transition in 1T-TiSe₂ has repeatedly been associated with an excitonic insulator instability^{12,13}. In this model, the transition is driven by the spontaneous formation of excitons, which can occur in semiconductors or semimetals when the bandgap or band overlap becomes smaller than the exciton binding energy. The exciton formation requires a low concentration of mobile charge carriers and a correspondingly poorly screened Coulomb interaction. In the case of 1T-TiSe₂, the narrow gap/overlap system would become unstable to the formation of Ti 3d – Se 4p excitons and would exhibit a new periodicity governed by the wave vector connecting the corresponding valence and conduction band pockets. This purely electronic process seems consistent with the measured ultrafast response times and a screening-based interpretation, as bound excitons would certainly be screened by the photo-injected carriers. For a transition temperature of 200 K, energy-time uncertainty yields a response time of 35 fs for such a purely excitonic process (see lower solid line in Fig. 3a). More generally, however, our results are in line with the

screening of any (unspecified) interaction between Se 4p and Ti 3d states described by an effective matrix element V_{eff}^2 . The spectral weight of the folded Se 4p state will scale with V_{eff}^2 and we can estimate the effect of screening within the Thomas-Fermi approach¹⁴. This yields $V_{\text{eff}}^2 \approx (1 + \text{const.} \times n^{1/3})$, in reasonable agreement with our experimental data (see Fig. 3b). A deeper analysis, particularly with respect to the puzzling story of the CDW phase transition in 1T-TiSe₂, requires a sophisticated theoretical description which, for instance, considers the quantum kinetics of screening in non-equilibrium systems or effects arising from the photo-doping of excitonic insulators.

We finally address the question of to what extent the screening by the nascent carriers affects the properties of 1T-TiSe₂ on the 20-fs timescale. As discussed above, the response of the atomic lattice is slow so that structurally the sample will still resemble the CDW phase on this timescale. However, the valence electronic structure becomes substantially modified as soon as screening becomes effective. In fact, electronically, the system undergoes a transition from a poorly conducting CDW state into a metallic phase within 20 fs. This is not only implied by the screening scenario but is also directly visible in Fig. 2d, which shows the transient metallization of the Ti 3d band only 30 fs after the optical excitation.

The ultrafast breakdown of long-range charge order that we report here is much faster than the material-characteristic oscillations of collective modes that are commonly thought to limit the response times in photoinduced processes. This surprising result may therefore stimulate

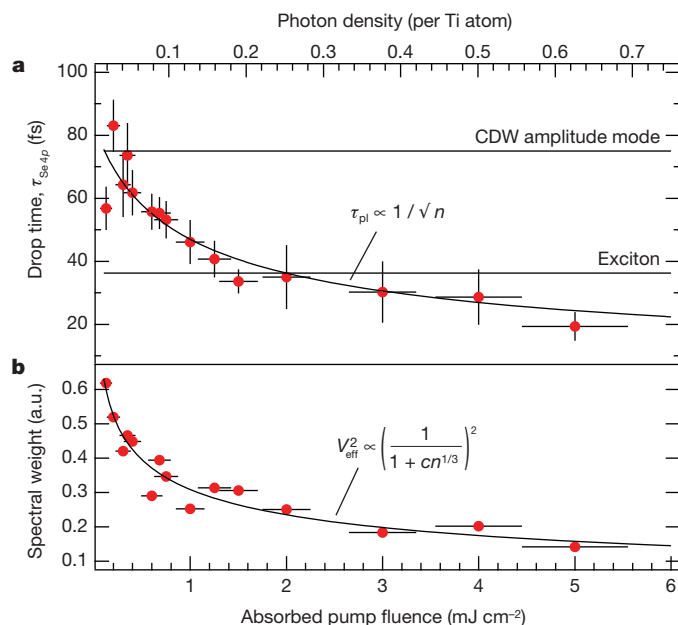


Figure 3 | Fluence dependence of the photoinduced transition. **a**, Signal drop time, $\tau_{\text{Se } 4p}$, of the folded Se 4p band (filled red circles) as a function of the absorbed pump fluence. The $\sim 10\%$ error bars in the absorbed pump fluence are determined from the uncertainty in the optical constants of 1T-TiSe₂, the stability of the pump laser, and the uncertainty in the pump pulse diameter at the sample position. The error bars in the drop time indicate the standard deviation of the fits to the data. **b**, Minimum transient spectral weight in the folded Se 4p band (red dots) as a function of absorbed fluence. Equation of curve fitted to data is shown in each panel. a.u., arbitrary units.

new concepts for ultrafast switching devices. Furthermore, in our study we have essentially monitored the intensity of a superlattice Bragg peak in electron momentum space. Thus, from a methodological point of view, the presented time-resolved ARPES approach complements present-day time-resolved diffraction experiments^{28,29}, with the advantage of an exceptionally high temporal resolution.

METHODS SUMMARY

Sample preparation. 1T-TiSe₂ single crystals were grown from the elements by chemical vapour transport using iodine as transport agent. Before the photoemission measurements, the samples were cleaved *in situ* at room temperature in ultrahigh vacuum.

Photoemission measurements. Static ARPES experiments were conducted at beamline 7.0.1 of the Advanced Light Source at Berkeley with a Scienta R4000 electron spectrometer. The photon energy was 119 eV and the overall energy resolution was ~ 30 meV. Femtosecond time-resolved ARPES measurements were conducted at the University of Kiel with a SPECS Phoibos 150 electron spectrometer. Here, the photon energy was 43 eV and the overall energy resolution in the experiment was ~ 400 meV. The light source for the time-resolved experiments was an argon-filled hollow-fibre waveguide (XUUS, KMLabs) for high harmonic generation operated with a 3 kHz Ti:sapphire amplifier system (Dragon, KMLabs, pumped by an Empower 30, Spectra Physics).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 6 September 2010; accepted 11 January 2011.

Published online 9 March 2011.

- Yonemitsu, K. & Nasu, K. Theory of photoinduced phase transitions in itinerant electron systems. *Phys. Rep.* **465**, 1–60 (2008).
- Bargheer, M., Zhavoronkov, N., Woerner, M. & Elsaesser, T. Recent progress in ultrafast X-ray diffraction. *ChemPhysChem* **7**, 783–792 (2006).
- Bigot, J.-Y., Vomer, M. & Beaupaire, E. Coherent ultrafast magnetism induced by femtosecond laser pulses. *Nature Phys.* **5**, 515–520 (2009).

- Chollet, M. *et al.* Gigantic photoresponse in $\frac{1}{4}$ -filled-band organic salt (EDO-TTF)₂PF₆. *Science* **307**, 86–89 (2005).
- Demsar, J., Biljaković, K. & Mihailović, D. Single particle and collective excitations in the one-dimensional charge density wave solid K_{0.3}MoO₃ probed in real time by femtosecond spectroscopy. *Phys. Rev. Lett.* **83**, 800–803 (1999).
- Rini, M. *et al.* Control of the electronic phase of a manganite by mode-selective vibrational excitation. *Nature* **449**, 72–74 (2007).
- Haight, R. & Silberman, J. A. Surface intervalley scattering on GaAs(110): direct observation with picosecond laser photoemission. *Phys. Rev. Lett.* **62**, 815–818 (1989).
- Fann, W. S., Storz, R., Tom, H. W. K. & Bokor, J. Electron thermalization in gold. *Phys. Rev. B* **46**, 13592–13595 (1992).
- Bauer, M. *et al.* Direct observation of surface chemistry using ultrafast soft-X-ray pulses. *Phys. Rev. Lett.* **87**, 025501 (2001).
- Perfetti, L. *et al.* Time evolution of the electronic structure of 1T-TaS₂ through the insulator-metal transition. *Phys. Rev. Lett.* **97**, 067402 (2006).
- Schmitt, F. *et al.* Transient electronic structure and melting of a charge density wave in TbTe₃. *Science* **321**, 1649–1652 (2008).
- Cercellier, H. *et al.* Evidence for an excitonic insulator phase in 1T-TiSe₂. *Phys. Rev. Lett.* **99**, 146403 (2007).
- Di Salvo, F. J., Moncton, D. E. & Waszczak, J. V. Electronic properties and superlattice formation in the semimetal TiSe₂. *Phys. Rev. B* **14**, 4321–4328 (1976).
- Ashcroft, N. W. & Mermin, N. D. *Solid State Physics* (Brooks/Cole, Belmont, 1976).
- Ohta, T., Bostwick, A., Seyller, T., Horn, K. & Rotenberg, E. Controlling the electronic structure of bilayer graphene. *Science* **313**, 951–954 (2006).
- Kondo, T., Khasanov, R., Takeuchi, T., Schmalian, J. & Kaminski, A. Competition between the pseudogap and superconductivity in the high-*T_c* copper oxides. *Nature* **457**, 296–300 (2009).
- Mazin, I. Superconductivity gets an iron boost. *Nature* **464**, 183–186 (2010).
- Rosnagel, K., Kipp, L. & Skibowski, M. Charge-density-wave phase transition in 1T-TiSe₂: excitonic insulator versus band-type Jahn-Teller mechanism. *Phys. Rev. B* **65**, 235101 (2002).
- Kidd, T. E., Miller, T., Chou, M. Y. & Chiang, T.-C. Electron-hole coupling and the charge density wave transition in TiSe₂. *Phys. Rev. Lett.* **88**, 226402 (2002).
- Rundquist, A. *et al.* Phase matching of soft-X-ray harmonic emission in hollow-core fibers. *Science* **280**, 1412–1415 (1998).
- Anisimov, S. I., Kapeliovich, B. L. & Perel'man, T. L. Electron-emission from surface of metals induced by ultrashort laser pulses. *Sov. Phys. JETP* **39**, 375–377 (1974).
- Cavalleri, A. *et al.* Evidence for a structurally-driven insulator-to-metal transition in VO₂: a view from the ultrafast timescale. *Phys. Rev. B* **70**, 161102(R) (2004).
- Wall, S. *et al.* Quantum interference between charge excitation paths in a solid-state Mott insulator. *Nature Phys.* advance online publication, doi:10.1038/nphys1831 (5 December 2010).
- Holy, J. A., Woo, K. C., Klein, M. V. & Brown, F. C. Raman and infrared studies of superlattice formation in TiSe₂. *Phys. Rev. B* **16**, 3628–3637 (1977).
- El Sayed, K., Schuster, S., Haug, H., Herzel, F. & Henneberger, K. Subpicosecond plasmon response: buildup of screening. *Phys. Rev. B* **49**, 7337–7344 (1994).
- Bányai, L., Vu, Q. T., Mieck, B. & Haug, H. Ultrafast quantum kinetics of time-dependent RPA-screened Coulomb scattering. *Phys. Rev. Lett.* **81**, 882–885 (1998).
- Huber, R. *et al.* How many-particle interactions develop after ultrafast excitation of an electron-hole plasma. *Nature* **414**, 286–289 (2001).
- Sokolowski-Tinten, K. *et al.* Femtosecond X-ray measurement of coherent lattice vibrations near the Lindemann stability limit. *Nature* **422**, 287–289 (2003).
- Siwick, B. J., Dwyer, J. R., Jordan, R. E. & Miller, R. J. D. An atomic-level view of melting using femtosecond electron diffraction. *Science* **302**, 1382–1385 (2003).
- Miaja-Avila, L. *et al.* Laser-assisted photoelectric effect from surfaces. *Phys. Rev. Lett.* **97**, 113604 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements M.B. and S.M. thank M. Aeschlimann for support and discussion. M.B. and S.M. also thank M. Murnane and H. Kapteyn for their support through the NSF EUV ERC. A.C. acknowledges support from the JILA Physics Frontier Center. This work was supported by the German Science Foundation (DFG) within the SFB 855 (C.S., M.B., L.K., K.R.) and by the European Community's FP7 under Marie Curie International Outgoing Fellowship GA 253316 (S.M.). Operation of the Advanced Light Source is supported by the US Department of Energy, Office of Basic Energy Sciences.

Author Contributions M.B. and K.R. conceived the experiment and wrote the paper. T.R., S.H., M.W., B.S., S.M., L.K., M.B. and K.R. realized the experimental time-resolved ARPES setup. A.C., L.M.A. and Y.L. designed and fabricated the EUV multilayer mirrors. T.R., S.H., M.W., C.S. and A.S. collected the time-resolved photoemission data and performed the data analysis. M.K. and K.R. collected and analysed the static photoemission data at the Advanced Light Source. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.B. (bauer@physik.uni-kiel.de; experimental technique) or K.R. (rosnagel@physik.uni-kiel.de; experimental data).

METHODS

Sample preparation and photoemission set-up. 1T-TiSe₂ single crystals were grown from the elements by chemical vapour transport using iodine as transport agent. Samples were mounted on a cryogenic manipulator and cleaved *in situ* under ultrahigh vacuum (UHV) conditions at a base pressure of 3×10^{-10} mbar. Conventional ARPES measurements were conducted at beamline 7.0.1 of the Advanced Light Source at Berkeley, with a Scienta R4000 electron spectrometer. The photon energy was 119 eV and the overall energy resolution was ~ 30 meV. Femtosecond time-resolved ARPES measurements were conducted at the University of Kiel. Photoemitted electrons were detected using a hemispherical electron energy analyzer (SPECS, Phoibos 150) equipped with a two-dimensional detection unit for parallel energy and momentum detection. The total energy resolution of the experiment was mainly governed by the spectral broadening of the femtosecond XUV pulses and was determined to be ~ 400 meV. An independent characterization of the high harmonic radiation with a grating spectrometer showed that the spectral width of the used 27th harmonic was 340 ± 40 meV. Typical integration times for analysis-grade spectra were 3 min. High-quality data as shown in Fig. 2 required an integration time of 15 min.

Pulsed XUV light source. The light source used for the pump–probe photoemission experiments was a 3-kHz Ti:sapphire amplifier system (Dragon, KMLabs, pumped by an Empower 30, Spectra Physics) delivering infrared pulses at 790 nm, 1.2 mJ pulse energy and 32 fs pulse duration. For photoemission, 80% of the pulse energy was used to generate high harmonic femtosecond XUV pulses in an argon-filled hollow-fibre waveguide (XUUS, KMLabs). A pair of multilayer mirrors (total reflectivity, 13%) selected the 27th harmonic ($h\nu = 43$ eV) out of the harmonic spectrum and focused it at an angle of 45° onto the sample mounted in the UHV system. The intensity of the 27th harmonic at the sample position was measured *in situ* using a calibrated XUV photodiode (SXUV 20 HS1, International Radiation Detections) yielding a fluence in the 10^9 photons s^{-1} regime. The residual 20% of the amplifier output was available for infrared photoexcitation of the 1T-TiSe₂ sample. The temporal pulse profile of the pump beam was characterized using the frequency-resolved optical gating technique. The pulse-width of the XUV pulses was estimated to a value < 10 fs from the LAPE infrared–XUV cross-correlation traces of the Se 4*p* band signal of the 1T-TiSe₂ sample³⁰ (Fig. 2f).

Hedgehog/Wnt feedback supports regenerative proliferation of epithelial stem cells in bladder

Kunyoo Shin¹, John Lee¹, Nini Guo², James Kim¹, Agnes Lim¹, Lishu Qu¹, Indira U. Mysorekar³ & Philip A. Beachy¹

Epithelial integrity in metazoan organs is maintained through the regulated proliferation and differentiation of organ-specific stem and progenitor cells. Although the epithelia of organs such as the intestine regenerate constantly and thus remain continuously proliferative¹, other organs, such as the mammalian urinary bladder, shift from near-quiescence to a highly proliferative state in response to epithelial injury^{2–4}. The cellular and molecular mechanisms underlying this injury-induced mode of regenerative response are poorly defined. Here we show in mice that the proliferative response to bacterial infection or chemical injury within the bladder is regulated by signal feedback between basal cells of the urothelium and the stromal cells that underlie them. We demonstrate that these basal cells include stem cells capable of regenerating all cell types within the urothelium, and are marked by expression of the secreted protein signal Sonic hedgehog (Shh). On injury, Shh expression in these basal cells increases and elicits increased stromal expression of Wnt protein signals, which in turn stimulate the proliferation of both urothelial and stromal cells. The heightened activity of this signal feedback circuit and the associated increase in cell proliferation appear to be required for restoration of urothelial function and, in the case of bacterial injury, may help clear and prevent further spread of infection. Our findings provide a conceptual framework for injury-induced epithelial regeneration in endodermal organs, and may provide a basis for understanding the roles of signalling pathways in cancer growth and metastasis.

The multi-layered bladder epithelium consists of a luminal layer of fully differentiated, usually binucleate umbrella cells⁵, which overlie intermediate cells with limited proliferative potential, and long-term label-retaining basal cells able to produce large colonies on culture *in vitro*⁶. These urothelial layers are separated by a basement membrane from the lamina propria, a thin layer of fibroblast-like stromal cells, and submucosal, smooth muscle and serous layers. Infections of the urinary tract, occurring in 10% of women annually⁷, are a common cause of injury to the bladder, and can be modelled in mice by infection with uropathogenic bacteria isolated from patients^{2,3,8}.

To establish baseline parameters of the bladder regenerative response, we induced injury in female mice by transurethral instillation of UTI89 (Supplementary Figs 1, 2a, 3a, b), a uropathogenic strain of *Escherichia coli*, or of the injurious compound protamine sulphate (PS). We found that expression of the proliferative marker Ki67 increased from near zero to 72% of epithelial and 28% of stromal cells within 24 h of UTI89 infection (Fig. 1a, b). The number of urothelial cell layers and total cells expressing the basal cell marker cytokeratin 5 (Ck5, also known as Krt5) was markedly expanded (Supplementary Fig. 3c, d), suggesting that injury-induced proliferation occurs primarily in basal urothelial cells. Instillation of increasing PS concentrations also induced Ki67 expression, reaching a plateau of ~35% at 20 mg ml⁻¹ PS and above (Fig. 1c and Supplementary Fig. 4); Ck5-positive cell number increased more modestly with PS injury. Interestingly, stromal cells showed no increased proliferation, even at PS concentrations that saturate the epithelial response (Fig. 1c).

Shh is expressed in and has a role in development of urogenital sinus derivatives, including bladder and prostate^{9–11}. In the adult urothelium, we found that Shh is expressed primarily in Ck5-positive basal cells, as indicated by immunostaining and GFP expression in a *Shh-eGFP* BAC transgenic strain (Fig. 1d and Supplementary Fig. 5a). Hedgehog (Hh) pathway activity, indicated by a reporter *Gli1-LacZ* strain¹² (Supplementary Fig. 6a, b), is restricted to stromal, submucosal and muscle layers, outside the urothelium (Fig. 1e), and depends on the Shh signal,

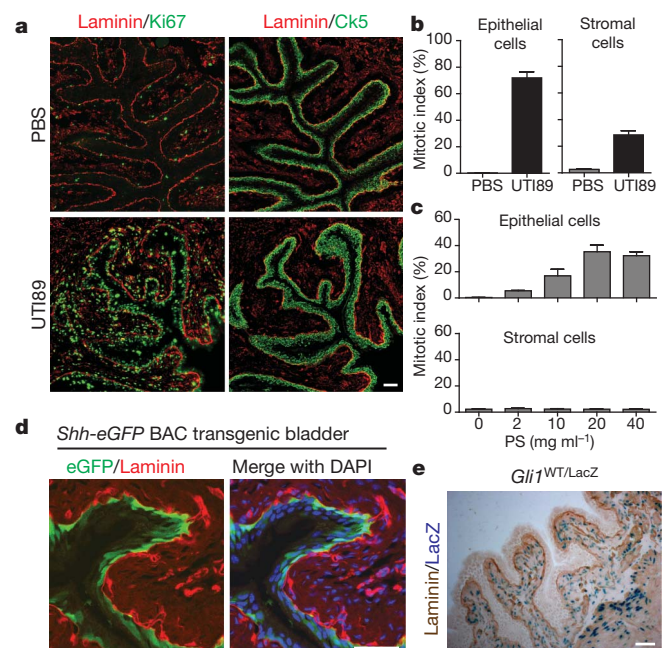


Figure 1 | Injury-induced proliferation and Hedgehog signalling in the bladder. **a**, UTI89 instillation induces proliferation of basal epithelial and stromal cells of the bladder. Ki67, Ck5 and laminin immunostaining highlight proliferation, basal epithelial cells and the basement membrane, respectively, in bladders 24 h after instillation of UTI89. Adjacent sections were 10 μ m apart. **b**, Quantification of epithelial and stromal cell proliferation in response to bacterial injury. Ki67-positive cells are shown as a per cent of total 4',6-diamidino-2-phenylindole (DAPI)-staining nuclei. **c**, Quantification of epithelial and stromal cell proliferation in response to chemical injury. Ki67-positive cells are shown as a per cent of total DAPI-staining nuclei 24 h after instillation of the indicated concentrations of PS. Note the absence of a proliferative response in the stroma. For panels **b** and **c**, data are from 3 bladders, 2 sections each, and are shown as mean \pm s.e.m.; numerical data are in Supplementary Tables 1 and 2, respectively. **d**, Expression of eGFP in basal epithelial cells from a *Shh-eGFP* BAC transgenic mouse. **e**, *Gli1-LacZ* expression in the stromal compartment. Bladder sections from *Gli1*^{LacZ/WT} mice were co-stained with X-gal and anti-laminin. Scale bars in panels **a**, **d** and **e** represent 50 μ m.

¹Department of Developmental Biology, Institute for Stem Cell Biology and Regenerative Medicine, Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305, USA. ²Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. ³Department of Obstetrics and Gynecology, Washington University School of Medicine, St. Louis, Missouri 63110, USA.

as indicated by reduced *Gli1* and *Ptc* (also known as *Ptch1*) expression on treatment with a Shh-blocking antibody (Supplementary Fig. 5b, c).

Shh and *Gli1* mRNA levels both increased in response to injury (Supplementary Fig. 5d and Supplementary Table 3), and Shh protein expression extended to multiple layers of Ki67-positive epithelial cells, including basal cells and Ck5-positive basal-like cells that result from injury-induced proliferation (Supplementary Fig. 5e). Shh response was augmented by injury, requiring only four instead of twenty hours of X-gal staining for detection in *Gli1-LacZ* reporter mice, but nevertheless remained confined to the stromal compartment (Supplementary Fig. 5f).

Basal cells have been suggested to function as stem cells in many epithelia including bladder and prostate^{6,13}; we marked *Shh*-expressing cells *in vivo* using a CreER tamoxifen-dependent site-specific recombinase expressed under the control of the *Shh* promoter (*Shh*^{CreER})¹⁴ in combination with *R26*^{mTmG}, a Cre-sensitive bi-fluorescent reporter¹⁵. In *Shh*^{CreER/WT}; *R26*^{mTmG/WT} mice, the membrane-associated tomato fluorescent protein (mT) is expressed until tamoxifen (TM) injection (Supplementary Figs 2b, 7a), after which membrane-associated GFP (mG) marks *Shh*-expressing Ck5-positive cells in basal urothelium (Supplementary Fig. 8a). With three rounds of bacterial injury and recovery after TM injection (Supplementary Fig. 2b), mG labelled all or most urothelial cells including Ck5-positive basal cells, intermediate cells, and Ck5-negative luminal umbrella cells marked by expression of uroplakin 3 (ref. 16; Fig. 2a and Supplementary Figs 7c, 8a), indicating multipotency of *Shh*-expressing cells. With reduced TM treatment and a single round of bacterial injury, we observed less extensive marking in isolated, vertically coherent patches that included basal, intermediate and luminal cells (data not shown); overall regeneration thus appears to result from the combined activation of many local urothelial units.

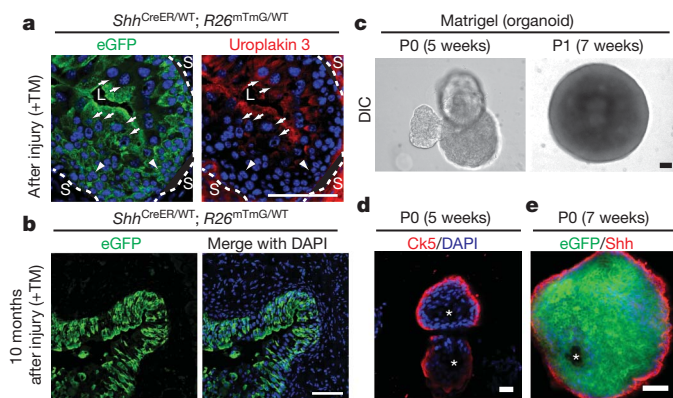


Figure 2 | Shh-expressing basal cells repopulate the urothelium and form organoids *in vitro*. **a**, eGFP-marked *Shh*-expressing cells generate luminal cells positive for uroplakin 3. *Shh*^{CreER/WT}; *R26*^{mTmG/WT} mice were treated with TM and subjected to three rounds of injury. Uroplakin-3-positive and negative cells are denoted by white arrows and arrowheads, respectively. Dotted lines demarcate urothelium and stroma (L, lumen; S, stroma). **b**, Long-term regenerative capacity of eGFP-marked *Shh*-expressing cells. After TM injection and seven rounds of injury over 10 months, mG expression in bladder from *Shh*^{CreER/WT}; *R26*^{mTmG/WT} mice marks most urothelial cells. See Supplementary Fig. 2b for experimental schemes. **c**, Extended culture of *Shh*-expressing cells in Matrigel. Single eGFP-positive bladder cells isolated from TM-injected *Shh*^{CreER/WT}; *R26*^{mTmG/WT} mice and cultured in Matrigel for 5 weeks formed organoids. Dissociated cells from primary culture (P0) organoids also generated new organoids on subsequent passage (P1) in Matrigel culture. DIC, differential interference contrast. **d**, Confocal analysis of a bladder organoid. The organoid has multiple layers of epithelial cells with Ck5-expressing cells in the outer layer that contacts the extracellular matrix, and inner cells that line a luminal space and do not express Ck5. **e**, A section through the wall of an organoid grown in Matrigel and immunostained for Shh. Note eGFP expression in all cells, indicative of *Shh* expression in the cell initially cultured, but loss of Shh immunostaining from cells that are not in the outer layer. Scale bars represent 50 μ m; asterisks denote organoid lumen.

We found similar labelling in the majority of the urothelium in mice carrying marked *Shh*-expressing cells and subjected to seven cycles of bacterial infection and recovery over a period of 10 months (Fig. 2b and Supplementary Fig. 8b). This persistence through lengthy and repeated periods of intense proliferation suggests that *Shh*-expressing cells have a capacity for self-renewal. Similarly marked cells traced through a 10-month period without injury produced less extensive labelling that nevertheless included Ck5-positive and Ck5-negative cells (Supplementary Fig. 8c, d), indicating that *Shh*-expressing cells also participate in regular homeostatic turnover in the absence of injury.

Similar experiments using *Gli1-CreER*¹⁷ to mark *Shh*-responsive stromal cells revealed that mG expression marks most cells of the lamina propria, and no urothelial cells (Supplementary Figs 7b, 9a, b). We also timed TM administration such that cells were marked during the proliferative response to injury, and noted no injury-induced plasticity with regard to segregation of *Shh* and *Gli1* expression within epithelial and stromal compartments, respectively (Supplementary Fig. 10a–c).

Enhanced GFP (eGFP)-positive cells isolated by fluorescence-activated cell sorting (FACS) from the bladders of TM-injected *Shh*^{CreER/WT}; *R26*^{mTmG/WT} mice (WT, wild type; Supplementary Figs 2c, 11a, b) formed spheres within 2 weeks of culture in suspension or in Matrigel (Supplementary Fig. 11a–d). Approximately 5–6% of isolated cells formed primary spheres in culture, with most of the remaining cells dying rapidly, probably as a result of stress resulting from the isolation procedure; about 30–40% of cells within primary spheres were able to form secondary spheres in subsequent cultures (Supplementary Fig. 11e, f). After 5–7 weeks of culture in Matrigel, single *Shh*-expressing cells formed cyst-like organoids 700–1,000 μ m in diameter that resemble the bladder in containing multiple layers of epithelial cells with Ck5- and *Shh*-expressing cells in the outer layer that contacts the extracellular matrix, and inner cells that line a luminal space and express neither Ck5 nor *Shh* (Fig. 2c–e, Supplementary Fig. 11g, i and Supplementary Movie 1). Single cells from these organoids were capable of self-renewing by generating new organoids in subsequent cultures (Fig. 2c, Supplementary Fig. 11h, j and Supplementary Movie 2). Our *in vivo* and *in vitro* evidence thus indicates that *Shh*-expressing basal urothelial cells include multipotent stem cells that are capable of self-renewal and differentiation.

The *Gli1* member of the Gli family of transcriptional effectors that mediate transcriptional response to Hh signalling (Supplementary Fig. 6a, b), although not essential for viability or fertility¹², can contribute significantly to pathway activity. For example, the incidence of medulloblastoma in *Ptc*^{+/-} mice is reduced ~10-fold on deletion of the *Gli1* gene¹⁸. We found that epithelial proliferation induced by UT189 instillation was nearly absent in *Gli1* mutant bladders at the 24–48 h peak of wild-type proliferation, with no additional layers of basal-like cells; a later peak in proliferation was ~1/2 of the wild-type maximum (Fig. 3a, b and Supplementary Fig. 12a). Proliferation of stromal cells was also affected in *Gli1* mutants, with a similar delay and decrease in Ki67 expression (Fig. 3a, b). Injection of a *Shh*-blocking antibody reduced expression of pathway targets *Gli1* and *Ptc* (Supplementary Fig. 5b, c), and correspondingly reduced proliferative responses in both epithelial and stromal compartments (Supplementary Fig. 12b–d and Supplementary Table 5). The response to chemical injury was also markedly reduced in *Gli1* mutants (Fig. 3c and Supplementary Fig. 13a, b), and these results together indicate that stromal *Gli1*-mediated response to the epithelial *Shh* signal promotes proliferative activity in response to both chemical and bacterial injury.

We tested the role of proliferation in restoring urothelial integrity by instilling fluorescein isothiocyanate (FITC)-conjugated dextran after UT189-mediated injury in wild-type and *Gli1* mutants. At 6 h after infection, by which time umbrella cells have exfoliated⁸, wild-type and mutant bladders both showed penetration of FITC-dextran into interstitial spaces of the urothelium (not shown); at 24 h, however,

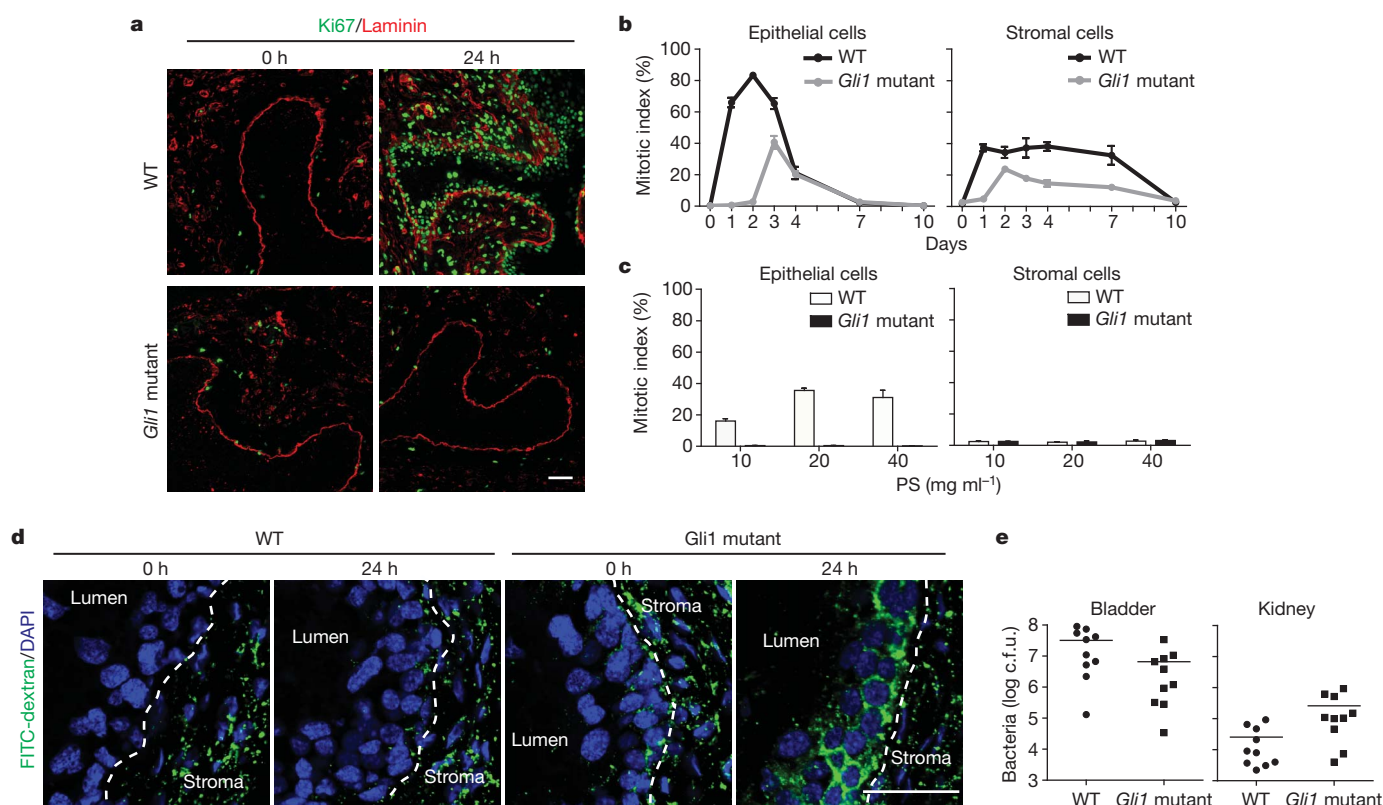


Figure 3 | *Gli1* mediates injury-induced proliferation, restoration of urothelial integrity and reduction of infectious spread. **a**, *Gli1* loss delays and attenuates proliferative response to bacterial injury. Bladders from wild-type or homozygous *Gli1* mutant mice were analysed at the indicated times after UTI89 instillation by immunostaining for Ki67 and laminin. **b**, Quantification of *Gli1* effect on epithelial and stromal cell proliferation. Ki67-positive cells are shown as a per cent of total DAPI-staining nuclei at the indicated times after UTI89 instillation in wild-type and *Gli1* mutant bladders. **c**, *Gli1* loss blocks the proliferative response of epithelial cells to chemical injury. Ki67-positive cells are shown as a per cent of total DAPI-staining nuclei 24 h after instillation of the indicated concentrations of PS. In panels **b** and **c**, data are shown as mean \pm s.e.m. from 3 bladders, 2 sections each, and numerical data are shown in Supplementary Tables 4 and 6, respectively. **d**, Paracellular permeability in

injured bladders from *Gli1* mutant mice. Bladders from wild-type and *Gli1* homozygotes were instilled with UTI89 and analysed at the times indicated after infection, with FITC-dextran instillation preceding bladder collection by 1.5 h. Dotted lines demarcate the border between urothelium and stroma. Note that normal reduced levels of paracellular permeability are restored by 24 h in wild-type but not *Gli1* homozygotes. **e**, Infectious spread to kidneys is enhanced by *Gli1* loss. Bacterial titres 24 h after UTI89 instillation were lower in bladders from *Gli1* homozygotes as compared to wild-type ($6.6 \pm 3.27 \times 10^6$ versus $3.2 \pm 1.03 \times 10^7$ colony-forming units (c.f.u.); $P < 0.05$). In contrast, bacterial titres were significantly higher in kidneys from *Gli1* homozygotes as compared to wild-type ($2.59 \pm 1.0 \times 10^5$ versus $2.55 \pm 1.0 \times 10^4$; $P < 0.05$). Data are presented as mean \pm s.e.m. (10 mice), and significance was calculated by an unpaired Student's *t*-test. Scale bars represent 50 μ m.

wild-type but not mutant bladders had re-established exclusion of FITC-dextran from extracellular spaces (Fig. 3d).

We also found that, although bacterial titres were somewhat lower in the bladders of infected *Gli1* mutant mice (Fig. 3e), the kidneys of mutant mice contained more than tenfold higher numbers of bacteria (Fig. 3e). These findings indicate that in addition to helping restore epithelial integrity, rapid proliferation of urothelial cells during normal regeneration may help reduce the risk of bacterial spread from the bladder to the kidneys, perhaps by competing for adhesive interactions that otherwise might aid in bacterial ascent via the ureters to the kidneys¹⁹.

The requirement for stromal *Gli1* in mediating proliferative response to epithelial injury suggested the possibility of *Shh*/*Gli1*-dependent transcription of secreted signals within the stroma. From gene expression profiles (data not shown) and quantitative polymerase chain reaction with reverse transcription (RT-PCR) with RNA from injured and uninjured wild-type or *Gli1* mutant bladders, we found that *Wnt2*, *Wnt4* and *Fgf16* showed significant injury- and *Gli1*-dependent responses (Supplementary Fig. 14a, b). With RNA isolated from stromal and epithelial compartments using laser capture microdissection (LCM; Supplementary Fig. 15a, b), we found that levels of *Wnt2*, *Wnt4* and *Fgf16* transcripts increased with injury only in the stroma (Fig. 4a).

For further analysis we isolated basal, intermediate and umbrella cell layers from uninjured wild-type or *Gli1* mutant bladders

(Supplementary Fig. 15c, d); as umbrella cells were absent from regenerating bladders owing to bacterially induced exfoliation, we isolated a single luminal layer of intermediate cells, a single layer of Ck5-positive basal cells and two layers of Ck5-positive basal-like cells (basal-like1 and basal-like2) from wild type (Fig. 4b and Supplementary Fig. 15e) or, from *Gli1* homozygous mutants, intermediate cells and a single layer of basal cells (Supplementary Fig. 15f). With RNA from these microdissected cell layers we found that transcription of *Axin2*, a universal indicator of Wnt signal response^{20,21}, increased markedly in stromal cells and in basal and basal-like1 cells, to a lesser extent in the second basal-like2 layer, and not at all in the intermediate cell layer (Fig. 4c). No increase in *Axin2* levels could be seen in stromal or urothelial layers of injured, *Gli1*-mutant bladders (Fig. 4c). Increased Wnt response thus occurs in stromal cells and in urothelial cell layers in closest proximity to the stromal source of Wnt signals.

Shh expression also increased in basal and basal-like cells (Fig. 4d and Supplementary Fig. 16), indicating stromal *Gli1*-mediated positive feedback on epithelial *Shh* expression. Notably, however, *Shh* expression in basal cells also increased somewhat in the *Gli1* mutant, indicating an injury-induced effect on *Shh* expression that does not require *Gli1*-mediated feedback from the stroma.

We tested pharmacological modulators of Wnt signalling, and found that indomethacin treatment reduced *Axin2* transcripts twofold in the bladder (Supplementary Fig. 17a), indicating a reduction in Wnt

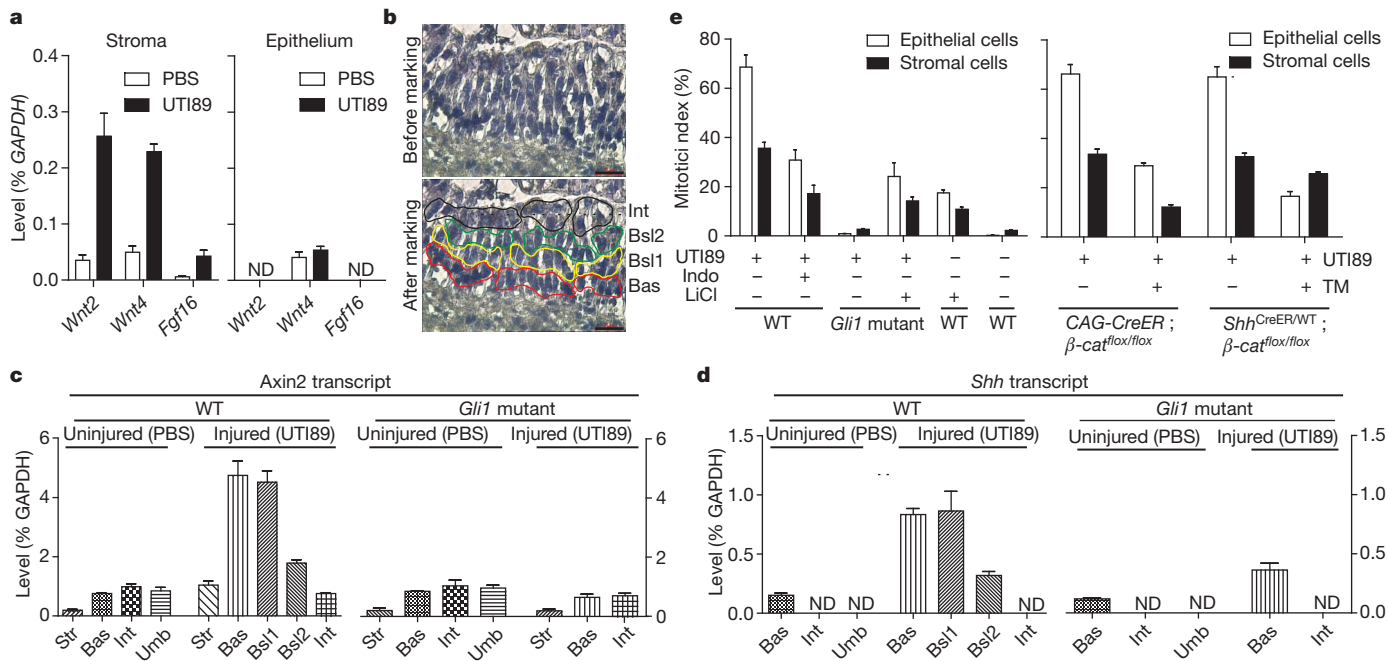


Figure 4 | Hedgehog-induced expression of stromal Wnt signals mediates urothelial and stromal proliferation. **a**, *Wnt2*, *Wnt4* and *Fgf16* expression in microdissected epithelium or stroma. Stromal expression of *Wnt2*, *Wnt4* and *Fgf16* increased significantly on injury. Although *Wnt4* RNA was detected in the epithelium, this expression did not increase on injury. ND, not detected. **b**, Laser capture microdissection of urothelial cell layers in regenerating bladder. Red, yellow, green and black lines illustrate selection and outlining of cells in basal (Bas), basal-like 1 (Bsl1), basal-like 2 (Bsl2), and intermediate (Int) layers, respectively, before microdissection. **c**, **d**, *Axin2* and *Shh* expression in microdissected cell layers from wild-type and *Gli1* mutant bladders 24 h after instillation of UTI89 or PBS. **c**, Expression of *Axin2* increased 5.5 fold in stroma ($P < 0.05$), and 6.3 fold in basal ($P < 0.05$), 5.9 fold in basal-like 1 ($P < 0.05$), and 2.3 fold in basal-like 2 ($P < 0.01$) as compared to uninjured basal cells. Expression of *Axin2* did not change significantly in basal and intermediate cells of injured *Gli1* homozygous mutants as compared to basal and intermediate

cells of uninjured *Gli1* mutant bladder. **d**, Expression of *Shh* in cells increased 5.5 fold in basal ($P < 0.01$), 5.7 fold in basal-like 1 ($P < 0.05$), and 2.1 fold in basal-like 2 ($P < 0.05$) as compared to uninjured basal cells. Expression of *Shh* in *Gli1* homozygous mutants increased 3 fold ($P < 0.05$) in injured as compared to uninjured basal epithelial cells. Data are presented as mean \pm s.e.m., and significance was calculated by a paired Student's *t*-test. Str, stroma; Umb, umbrella cells. **e**, Modulation of Wnt pathway activity in regenerative response to bacterial injury. Left, pharmacological reduction (indomethacin (Indo)) or augmentation (LiCl) of the Wnt signal response respectively decreases or increases bladder proliferation, with or without bacterial injury, as shown, in mice of the indicated genotype (see also Supplementary Fig. 17b, c, d). Right, tamoxifen-induced inactivation of β -catenin decreases proliferation in epithelium and stroma (*CAG-CreER*) or in epithelium (*Shh^{CreER}*). Data are presented as mean \pm s.e.m. from 3 bladders, 2 sections each (see also Supplementary Fig. 18e, f).

response²⁰, and correspondingly suppressed UTI89-induced proliferation in the epithelium and stroma (Fig. 4e and Supplementary Fig. 17b). We also found that LiCl²² increased *Axin2* transcripts (Supplementary Fig. 17a), and correspondingly induced proliferation in the epithelium and stroma of uninjured wild-type mice or substantially rescued the proliferative response of *Gli1* mutant mice (Fig. 4e and Supplementary Fig. 17c, d). We also observed a marked enhancement of the proliferative response to instillation of 2 mg ml⁻¹ PS, from ~5% to 30% in the epithelium and from 0 to ~15% in the stroma (Supplementary Fig. 18a, b), in mice heterozygous for the *Apc*^{min} mutation, which show constitutive Wnt pathway activity due to reduced dosage of a negative Wnt response regulator²³ (Supplementary Fig. 18c).

We inactivated the Wnt response with a homozygous conditional allele of the essential Wnt pathway component β -catenin (also known as Ctnnb1), and found that ablation in TM-injected mice by the ubiquitously expressed CAG-CreER reduced *Axin2* expression in injured bladder (Supplementary Fig. 18d) and correspondingly reduced injury-induced proliferation in both epithelial and stromal compartments (Fig. 4e and Supplementary Fig. 18e). In contrast, the proliferation defect produced in *Shh*^{CreER/WT}; β -catenin^{lox/lox} mice was restricted to basal epithelium (Fig. 4e and Supplementary Fig. 18f). Our pharmacological and genetic data indicate a role for Wnt pathway activity within basal epithelial and stromal cells in the activation of proliferative responses, with the potential role and importance of Fgf or other Hh-induced stromal signals remaining to be explored.

Our findings, summarized schematically in Supplementary Fig. 19, reveal an essential contribution by Hh and Wnt signals acting across the

epithelial–stromal boundary during bladder regeneration, and are reminiscent of the well-studied *Drosophila* embryonic segment, in which Hh and Wingless signalling across the parasegment boundary are essential for segmental patterning during development. Interestingly, however, Hh/Wnt feedback signalling in the *Drosophila* segment operates to specify future pattern within an undifferentiated epithelium, whereas in the bladder it functions to maintain differentiated structures in a mature organ.

Surprisingly, despite its dispensability for normal development, *Gli1* contributes significantly to bladder regeneration, thus providing a useful tool to demonstrate a role of regenerative proliferation not only in restoring urothelial integrity but also in preventing bacterial spread to the kidneys. Further studies will be required to determine whether Gli1 also contributes to the regeneration of other organs in which Hh signalling has a role^{11,24}; such flexibility in Gli1 transcriptional output would be consistent with the apparent flexibility of Gli1 in mediating qualitatively distinct responses to bacterial and chemical bladder injury, which differ markedly in the presence or absence of stromal cell proliferation.

As many as 10% of women experience infections of the urinary tract in a year^{7,25}, including cystitis and pyelonephritis, and ~26% of urinary tract infections recur within six months²⁶. Recent work has demonstrated that intracellular reservoirs of bacteria can form in umbrella cells or in transitional cells³, and exfoliation of infected cells harbouring intracellular bacterial reservoirs induced by PS or other agents has been suggested as potentially beneficial in clearing these infections. Our current data suggest that it may also be worth exploring the effect of

more directly activating the Wnt signalling pathway by treating with LiCl; on the other hand use of drugs such as indomethacin, which inhibit the regenerative response in the urothelium, may be contraindicated.

Another clinical arena in which our findings may be relevant is the growth and dissemination of cancers in which Hh ligand production in primary cells of the tumour triggers pathway activity in tumour stroma, which then expands and supports the growth of primary cells within the tumour^{27,28}. This tumour–stromal interaction might now be viewed as the growth-enhancing activation of a feedback circuit normally triggered by injury. Tumour–stromal interactions are also critical in metastasis, and the preferential colonization of particular tissues by individual tumour types might be determined by the competence of target tissues to respond to eliciting signals from the tumour cells, similar to the epithelial–stromal interaction noted here in bladder regeneration. In this connection it is interesting to note that tumours of the prostate, which like the bladder is a derivative of the urogenital sinus and also requires Hh pathway activity for its regeneration¹¹, metastasize most commonly to bone, lung and liver²⁹, a preference very similar to that of bladder tumours, which metastasize most commonly to liver, lung and bone³⁰. Further work will be required to identify regenerative signals in other organs and to determine how important their activities may be in cancer growth and metastasis.

METHODS SUMMARY

For bladder injury with chemical agents or bacteria, transurethral instillation was performed as described². For *in vivo* lineage tracing, transgenes expressing tamoxifen-activated Cre combined with the *R26^{mTmG}* bi-fluorescent reporter were used to heritably mark cells and their progeny. Transgenes expressing tamoxifen-activated Cre also were used to selectively inactivate a conditional allele of β -catenin. For bladder sphere and organoid culture, GFP-marked Shh-expressing basal cells from TM-treated *Shh^{CreER/WT}*, *R26^{mTmG/WT}* mice were FACS-sorted and cultured in suspension or in Matrigel. Quantitative RT–PCR was used for measurement of RNA levels in samples isolated from the entire bladder, or in material isolated by laser capture microdissection from the stromal or the urothelial compartments, or from individual cell layers within the urothelium.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 September 2010; accepted 17 January 2011.

Published online 9 March 2011.

- van der Flier, L. G. & Clevers, H. Stem cells, self-renewal, and differentiation in the intestinal epithelium. *Annu. Rev. Physiol.* **71**, 241–260 (2009).
- Hung, C. S., Dodson, K. W. & Hultgren, S. J. A murine model of urinary tract infection. *Nature Protocols* **4**, 1230–1243 (2009).
- Mysorekar, I. U. & Hultgren, S. J. Mechanisms of uropathogenic *Escherichia coli* persistence and eradication from the urinary tract. *Proc. Natl Acad. Sci. USA* **103**, 14170–14175 (2006).
- Mysorekar, I. U., Isaacson-Schmid, M., Walker, J. N., Mills, J. C. & Hultgren, S. J. Bone morphogenetic protein 4 signaling regulates epithelial renewal in the urinary tract in response to uropathogenic infection. *Cell Host Microbe* **5**, 463–475 (2009).
- Hicks, R. M. The mammalian urinary bladder: an accommodating organ. *Biol. Rev. Camb. Philos. Soc.* **50**, 215–246 (1975).
- Kurzrock, E. A., Lieu, D. K., Degraffenried, L. A., Chan, C. W. & Isseroff, R. R. Label-retaining cells of the bladder: candidate urothelial stem cells. *Am. J. Physiol. Renal Physiol.* **294**, F1415–F1421 (2008).
- Hooton, T. M. & Stamm, W. E. Diagnosis and treatment of uncomplicated urinary tract infection. *Infect. Dis. Clin. North Am.* **11**, 551–581 (1997).
- Klumpp, D. J. *et al.* Uropathogenic *Escherichia coli* induces extrinsic and intrinsic cascades to initiate urothelial apoptosis. *Infect. Immun.* **74**, 5106–5113 (2006).

- Podlasek, C. A., Barnett, D. H., Clemens, J. Q., Bak, P. M. & Bushman, W. Prostate development requires Sonic hedgehog expressed by the urogenital sinus epithelium. *Dev. Biol.* **209**, 28–39 (1999).
- Haraguchi, R. *et al.* Molecular analysis of coordinated bladder and urogenital organ formation by Hedgehog signaling. *Development* **134**, 525–533 (2007).
- Karhadkar, S. S. *et al.* Hedgehog signalling in prostate regeneration, neoplasia and metastasis. *Nature* **431**, 707–712 (2004).
- Bai, C. B., Auerbach, W., Lee, J. S., Stephen, D. & Joyner, A. L. Gli2, but not Gli1, is required for initial Shh signaling and ectopic activation of the Shh pathway. *Development* **129**, 4753–4761 (2002).
- Lawson, D. A., Xin, L., Lukacs, R. U., Cheng, D. & Witte, O. N. Isolation and functional characterization of murine prostate stem cells. *Proc. Natl Acad. Sci. USA* **104**, 181–186 (2007).
- Harfe, B. D. *et al.* Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. *Cell* **118**, 517–528 (2004).
- Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **45**, 593–605 (2007).
- Wu, X. R. & Sun, T. T. Molecular cloning of a 47 kDa tissue-specific and differentiation-dependent urothelial cell surface glycoprotein. *J. Cell Sci.* **106**, 31–43 (1993).
- Ahn, S. & Joyner, A. L. Dynamic changes in the response of cells to positive Hedgehog signaling during mouse limb patterning. *Cell* **118**, 505–516 (2004).
- Kimura, H., Stephen, D., Joyner, A. & Curran, T. Gli1 is important for medulloblastoma formation in *Ptc1^{+/-}* mice. *Oncogene* **24**, 4026–4036 (2005).
- Hagberg, L. *et al.* Ascending, unobstructed urinary tract infection in mice caused by pyelonephritogenic *Escherichia coli* of human origin. *Infect. Immun.* **40**, 273–283 (1983).
- Goessling, W. *et al.* Genetic interaction of PGE2 and Wnt signaling regulates developmental specification of stem cells and regeneration. *Cell* **136**, 1136–1147 (2009).
- Lustig, B. *et al.* Negative feedback loop of Wnt signaling through upregulation of conductin/axin2 in colorectal and liver tumors. *Mol. Cell. Biol.* **22**, 1184–1193 (2002).
- Klein, P. S. & Melton, D. A. A molecular mechanism for the effect of lithium on development. *Proc. Natl Acad. Sci. USA* **93**, 8455–8459 (1996).
- Su, L. K. *et al.* Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. *Science* **256**, 668–670 (1992).
- Fendrich, V. *et al.* Hedgehog signaling is required for effective regeneration of exocrine pancreas. *Gastroenterology* **135**, 621–631 (2008).
- Nicolle, L. E. Uncomplicated urinary tract infection in adults including uncomplicated pyelonephritis. *Urol. Clin. North Am.* **35**, 1–12 (2008).
- Foxman, B. Recurring urinary tract infection: incidence and risk factors. *Am. J. Public Health* **80**, 331–333 (1990).
- Yauch, R. L. *et al.* A paracrine requirement for hedgehog signalling in cancer. *Nature* **455**, 406–410 (2008).
- Tian, H. *et al.* Hedgehog signaling is restricted to the stromal compartment during pancreatic carcinogenesis. *Proc. Natl Acad. Sci. USA* **106**, 4254–4259 (2009).
- Bubendorf, L. *et al.* Metastatic patterns of prostate cancer: an autopsy study of 1,589 patients. *Hum. Pathol.* **31**, 578–583 (2000).
- Wallmroth, A. *et al.* Patterns of metastasis in muscle-invasive bladder cancer (pT2–4): an autopsy study on 367 patients. *Urol. Int.* **62**, 69–75 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Oro and J. Brooks for their critical reading of the manuscript, and the Stanford Center for Digestive Diseases for help with laser capture microdissection. This research was supported in part by grants from the Department of Defense and from the National Institutes of Health (P.A.B.) and a Pathway to Independence Award (K99/R00) to I.U.M. P.A.B. is an investigator of the Howard Hughes Medical Institute.

Author Contributions K.S. and P.A.B. conceived ideas and experimental design. K.S. performed the experiments. N.G. aided in immunohistochemical analysis, J.L. and J.K. helped with mouse strains, A.L. assisted with *in vitro* cell culture studies, L.Q. performed the genotyping of experimental mice, and I.U.M. helped analyse data. K.S. and P.A.B. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.A.B. (pbeachy@stanford.edu) or K.S. (kunyoos@stanford.edu).

METHODS

Mice. *Gli1*^{LacZ/WT} heterozygotes¹² from our colony were interbred to generate *Gli1*^{LacZ/LacZ} homozygotes. Heterozygous and wild-type littermates were used as controls. Female mice between 6 and 10 weeks of age were used for all injury experiments. *Shh-GFP* BAC strain was obtained from the GENSAT project at Rockefeller University. For lineage tracing experiments, *Shh*^{CreER/WT}, or *Gli1*^{CreER/WT} mice^{14,17} were crossed with *R26*^{mTmG/mTmG} strain¹⁵ to obtain *Shh*^{CreER/WT}; *R26*^{mTmG/WT}, or *Gli1*^{CreER/WT}; *R26*^{mTmG/WT}. *Shh*^{CreER/WT} or *CAG-CreER* were crossed with *β-catenin*^{flx/flx} mice to obtain *Shh*^{CreER/WT}; *β-catenin*^{flx/flx} or *CAG-CreER*; *β-catenin*^{flx/flx}. All mouse strains except as otherwise indicated were obtained from Jackson Laboratories. All bladder instillation procedures were performed under isoflurane anaesthesia, which was administered in a fume hood with a standard vaporizer (J. B. Baulch and Associates). All procedures were performed under a protocol approved by the Administrative Panel on Laboratory Animal Care at Stanford University.

Bacterial and chemical injury. For bacterial injury, a uropathogenic *E. coli* (UPEC) strain, UTI89, was grown for 16 h in a static culture, and inoculated via transurethral instillation of anaesthetized female mice at a concentration of 10^7 c.f.u. in 50 µl as previously described². Animals were maintained after infection for the indicated period of time or urine was collected after 4 h to confirm infection. For chemical injury, 50 µl of PS (Sigma) solution in PBS at 2 mg ml⁻¹, 10 mg ml⁻¹, 20 mg ml⁻¹ or 40 mg ml⁻¹ was delivered transurethally as indicated. Bladders were collected and analysed at distinct time-points after instillation.

Lineage tracing studies. For fate mapping of *Shh*-expressing cells before bacterial injury, *Shh*^{CreER/WT}; *R26*^{mTmG/WT} or *Gli1*^{CreER/WT}; *R26*^{mTmG/WT} mouse strains were injected intraperitoneally with 4 mg of TM (per 30 g body weight) daily for three consecutive days. Seven days after the last TM injection, mice were subjected to three transurethral instillations of UTI89 with 10-day intervals after the first and second infections, and 15 days after the third. Mice were killed and bladders then dissected for further analysis. For labelling of *Shh*- or *Gli1*-expressing cells before injury, mice were injected with TM (4 mg per 30 g body weight) for 3 days and analysed 5 days after the last injection. For lineage tracing of *Shh* or *Gli1* expressing cells during injury, *Shh*^{CreER/WT}; *R26*^{mTmG/WT} or *Gli1*^{CreER/WT}; *R26*^{mTmG/WT} mice were injected intraperitoneally with 4 mg of TM (per 30 g body weight) for 5 consecutive days, starting 2 days before infections to allow enough time for tamoxifen to be absorbed by the bladder. Six days after the last injection of TM, the entire procedure was repeated. Tissue sections were prepared 6 days after final injection of TM for further analysis. For long-term lineage tracing, *Shh*^{CreER/WT}; *R26*^{mTmG/WT} mouse strains were injected intraperitoneally with 4 mg of TM (per 30 g body weight) daily for 3 consecutive days. Seven days after the last TM injection, mice were subjected to seven transurethral instillations of UTI89, twice a month for 2 months and once a month for next 3 months. Five months after the last instillation, mice were killed and bladders then dissected for further analysis. For long-term lineage tracing experiments to study homeostasis, *Shh*^{CreER/WT}; *R26*^{mTmG/WT} mouse strains were injected with 4 mg of TM (per 30 g body weight) daily for 3 consecutive days, and bladders were analysed 10 months after the last TM injection.

In vitro culture of *Shh*-expressing cells. To isolate *Shh*-expressing cells, *Shh*^{CreER/WT}; *R26*^{mTmG/WT} mouse strains were injected intraperitoneally with 4 mg of TM (per 30 g body weight) daily for 3 consecutive days. Three days after TM injection, bladders were collected, inverted and inflated as described previously³¹. Inverted bladders were incubated in 0.25% Trypsin-EDTA containing 500 U ml⁻¹ collagenase for 2 h at 37 °C. Tissues were then minced and, after the lysis of red blood cells, a single-cell suspension was obtained by 10 min of trituration, followed by filtration through 40-µm cell strainers. Cells were sorted using a FACSAriaII cytometer (BD Biosciences), and analysis of flow cytometry data was performed using FlowJo Software (Treestar). Sorted cells were cultured in Ultra Low attachment plates (Corning) for suspension culture in media containing 1:1 mixture of conditioned media and DMEM (Invitrogen) supplemented with 50 ng ml⁻¹ EGF (PeproTech). Conditioned medium was prepared by growing 90% confluent V79 lung fibroblast-derived cells (ATCC) in DMEM with 5% FBS for 24 h. For Matrigel culture, a single-cell suspension was mixed with 200 µl of ice-cold Matrigel (reduced growth factors; BD Bioscience), layered onto 12-mm Transwell clear filters, and allowed to solidify at 37 °C. Pre-warmed growth medium was added above and below the gel and cells were grown until the time of analysis. For passing organoids in Matrigel culture, organoids were released from Matrigel matrix by depolymerizing Matrigel using Matrisphere Cell Recovery Solution (BD Bioscience) according to the manufacturer's instructions. Organoids were then dissociated into single cells by incubating with 0.25% Trypsin-EDTA for 30 min, followed by 5 min of trituration. Cells were then seeded in Matrigel as described earlier.

Antibody injection. Mice were injected intraperitoneally with either anti-mouse *Shh* antibody or isotype control (5E1 and 6B3, respectively; Developmental Studies Hybridoma Bank) daily for 7 days. Mice received 10 mg kg⁻¹ body weight of antibody for the first 3 days and 5 mg kg⁻¹ body weight thereafter. Bladders were dissected on the last day of injection to isolate RNA for qRT-PCR analysis. For assays of proliferation, mice were injected with 10 mg antibody per kg body weight daily for 3 days, and 5 mg antibody per kg body weight thereafter. Mice were infected with UTI89 on the fourth day of antibody injection.

Microscopy and laser capture microdissection. All images were obtained using a Zeiss LSM 510 inverted confocal microscope and prepared for publication with Zeiss LSM 5 Image Browser software and Adobe Photoshop CS3. Three-dimensional reconstructions of confocal images were generated using Imaris software (Bitplane Scientific Software). For LCM, bladder sections were prepared using an LCM staining kit (Ambion) and a Leica LMD6000 Laser Microdissection Microscope.

Quantitative RT-PCR. For qRT-PCR, bladders were frozen in liquid nitrogen and total RNAs isolated from frozen bladder tissue using RNeasy Plus Mini (Qiagen). qRT-PCR was performed using iScript one-step RT-PCR kit with SYBR Green and the Bio-Rad iCycler (BioRad). Assays were performed on bladders from six animals, and all values normalized to the GAPDH internal control, which does not vary on injury (data not shown). For material isolated by LCM, total RNA was prepared using RNAqueous-Micro RNA isolation kit (Ambion).

Indomethacin and LiCl treatment. Mice were injected with 2.5 mg kg⁻¹ of body weight with indomethacin or DMSO vehicle control every 12 h throughout the experiment²⁰. UTI89 was instilled 36 h after initial indomethacin injection, and bladders were analysed on the third day, 60 h after the initial indomethacin injection. For LiCl treatment, wild-type or *Gli1*^{LacZ/LacZ} mice received either 200 mg kg⁻¹ of body weight of LiCl in 100 µl of deionized water or 100 µl of deionized water as a vehicle control every day for 3 days by oral gavage. At the time of oral gavage, mice were also injected transurethally with 40 mg kg⁻¹ of body weight of LiCl. UTI89 infection was performed at 30 h after initial LiCl treatment, and bladders were analysed on the third day, 54 h after initial LiCl treatment.

Conditional ablation of *β-catenin*. *Shh*^{CreER/WT}; *β-catenin*^{flx/flx} or *CAG-CreER*; *β-catenin*^{flx/flx} female mice were injected intraperitoneally with 4 mg of TM (per 30 g body weight) daily for 3 consecutive days. For *CAG-CreER*; *β-catenin*^{flx/flx}, bladder was inoculated with UTI89 the day after the last injection of TM, and bladders were collected 24 h after infection. For *Shh*^{CreER/WT}; *β-catenin*^{flx/flx}, bladder was inoculated with UTI89 3 days after the last injection of TM, and collected 24 h after infection.

In vivo permeability assay. 10 mg ml⁻¹ of FITC-dextran (10000MW, Invitrogen) in a 50 ml volume of PBS³² was injected transurethally into the bladder lumen 1.5 h before collection of bladders at the indicated time points. Bladder sections were made and analysed for FITC.

Bacterial titration and analysis of urine. Urine from infected bladders was collected 6 h after UTI89 infection, and slides were prepared and stained for analysis using Cytospin and Hema3 staining kit (Fisher). Kidneys and bladders were dissected from wild-type littermate controls or *Gli1*^{LacZ/LacZ} mice 24 h after infection. Tissues were homogenized in 1 ml of PBS, and bacterial titres were determined by microtitre-plate dilution on LB plates².

X-gal histochemistry and immunofluorescence analysis. Bladders were dissected and embedded in OCT compound for snap freezing (Tissue-Tek). Frozen blocks were sectioned at 10-µm intervals using a Microm cryostat. For X-gal staining, frozen sections were fixed in 0.2% glutaraldehyde in PBS containing 5 mM EGTA and 2 mM MgCl₂ for 30 min at 4 °C. After washing twice with PBS containing 2 mM MgCl₂, sections were incubated with 1 mg ml⁻¹ of X-gal solution in PBS containing 0.02% NP40, 0.01% deoxycholic acid, 2 mM MgCl₂, 5 mM EGTA, 5 mM C₆FeK₃N₆, 5 mM C₆FeK₄N₆ for 4 h to overnight. Stained sections were counterstained with eosin solution (Sigma). For immunostaining, frozen tissue sections were fixed in 4% of paraformaldehyde for 30 min at 4 °C. After washing three times with PBS, tissue sections were blocked in 2% goat serum in PBS containing 0.25% Triton X-100 for 1 h, incubated with the following primary antibodies diluted in blocking solution overnight at 4 °C in a humidified chamber: rat anti-*Shh* (R&D, 1:200); rabbit anti-Ki67 (Abcam, 1:500); rabbit anti-Ck5 (Abcam, 1:500); chicken anti-laminin (Abcam, 1:300); mouse anti-uropilin 3 (Fitzgerald). Sections were washed three times with PBS containing 0.25% Triton X-100, incubated with DAPI and appropriate Alexa fluoro 488, 594, or 633 conjugated secondary antibodies diluted 1:1,000 in blocking solution for 2 h at 22 °C, washed again three times, and mounted on slides with Prolong Gold mounting reagent (Invitrogen). For immunostaining of organoids, Matrigel plugs were removed from the filter supports of Transwell plates, washed with PBS, and incubated with Matrisphere Cell Recovery Solution (BD Bioscience) for 20 min to partially dissolve the Matrigel. Subsequently, the gel-embedded bladder organoids were fixed in 4% paraformaldehyde for 15 min at 4 °C, followed by washing three times with PBS. Fixed organoids were then permeabilized in 0.25% Triton X-100 for 30 min, followed by incubation in blocking solution of 2% goat serum in

PBS containing 0.25% Triton X-100 for 1 h. Organoids were then incubated with primary antibodies diluted in blocking solution for 2 days at 4 °C in a humidified chamber. After three washes, bladder organoids were incubated with secondary antibody overnight in 4 °C, followed by three washes and mounting on Coverwell chamber (Grace Bio-Lab) with Prolong Gold mounting reagent (Invitrogen).

Statistical analysis. Statistical analysis was performed using GraphPad Prism software v.5. All data are presented as mean \pm s.e.m., and two group comparisons

were done with a two-tailed Student's *t*-test. A value of $P < 0.05$ was taken as statistically significant.

31. Kurzrock, E. A., Lieu, D. K., deGraffenried, L. A. & Isseroff, R. R. Rat urothelium: improved techniques for serial cultivation, expansion, freezing and reconstitution onto acellular matrix. *J. Urol.* **173**, 281–285 (2005).
32. Ibla, J. C. & Khoury, J. Methods to assess tissue permeability. *Methods Mol. Biol.* **341**, 111–117 (2006).

CAREERS

TURNING POINT Atmospheric scientist hopes to spread word on climate change **p.257**

BRIEF Proportion of female academic life scientists rises in United States **p.257**

NATUREJOBS For the latest career listings and advice www.naturejobs.com



FINANCE

Quantifiable prospects

Despite market gyrations, banks can offer mathematicians and physicists a way to put their acumen to lucrative use.

BY DAVID LINDLEY

After a couple of postdoctoral fellowships in the mid-1990s, Ben-Ami Gradwohl, then a physicist in Los Angeles, California, faced a tough market for tenure-track academic positions. And he had doubts about devoting his career to one particular strand of academic research. By chance, he started

reading a book on the mathematical methods behind option pricing — how to determine the costs of tradeable contracts for buying or selling assets — and came across references to some papers not generally available in libraries. He got in touch with a manager at Leland O'Brien Rubinstein (LOR), a now-defunct investment firm in Los Angeles, who offered to send him copies of the papers in return for

his résumé. That was when Gradwohl's career took a turn.

He did some consulting at LOR, which helped him to learn the mysteries of finance. Even after that, he was still considering pursuing research in high-tech industry or solar energy. But he took the plunge and went into finance full time, first with LOR and then with Nicholas-Applegate Capital Management in San Diego, California, where he analysed and guided investment portfolios. More than 15 years later, Gradwohl is still in the field, working as head of portfolio strategies and enterprise analytics at CitiMortgage, part of Citigroup in New York. It has been a long time since he did any programming or mathematical analysis, but his skill with numbers remains an asset.

Knowledge of sophisticated mathematical modelling techniques still provides a toehold for entry into some specialized jobs, but it's no longer quite so easy for people with a PhD and no prior knowledge of the field to walk into a well-paid job. Among scientists, there's a "widening pool of people who have realized that finance is attractive", says William Perraudin, director of Risk Control, a London-based company that supplies credit-risk software and models to banks and financial institutions. That growth in the talent pool has created greater competition for employment and, says Perraudin, "a gradual ratcheting up of the requirements for entry" in terms of experience and knowledge. On the other hand, recruitment is beginning to pick up again after the economic upheavals that began with the collapse of the Lehman Brothers global financial-services firm in 2008. Employment "hasn't yet recovered to pre-crisis levels", says Dara Lubarsky, a recruiter with the international agency Huxley Associates in New York. "But there's definitely an upswing."

MONEY AND MATHS

The incursion of deep mathematics into the financial world has its roots in a paper by economists Fischer Black and Myron Scholes (*J. Polit. Econ.* 81, 637–654; 1973). The perceived value of an option depends on the market movements that traders consider probable; Black and Scholes constructed a model that determines an optimum price using a simplified picture of typical stock-market variations. Their innovation led to 'hedged' investments or portfolios, which are designed to maintain their value regardless of market movements. It also offers insights into investment risks, allowing analysts to form quantitative ►

► estimates as to whether prices are likely to deviate from optimum values. From the model sprang a huge sector of finance based on the purchase and sale of derivatives: financial instruments whose value is derived from movements in the prices of the assets on which they are based.

The Black–Scholes model is a set of partial differential equations, so institutions interested in exploiting it need people at ease with mathematical systems and programming techniques. Jennifer Hodgdon had done postdoctoral work on models and software for materials science at Bell Laboratories in Murray Hill, New Jersey, before she went to work at Goldman Sachs in New York in 1994. Her job there, writing software to implement mathematical models for complex derivatives in commodities and foreign exchange, was in technical terms not so different from her previous work. She says she basically “didn’t know anything” about finance, but “picked it up quickly”.

When Hodgdon entered finance, she says, there was just one book — *Options, Futures, and Other Derivative Securities* by John Hull (Prentice Hall, 1988), later renamed *Options, Futures and Other Derivatives* and now in its eighth edition — that she and all other aspiring quantitative analysts read. But starting about 10 years ago, several universities, most of them in the United States and the United Kingdom, have developed master’s-level courses in ‘computational finance’ or ‘financial engineering’. These are typically aimed at students with an undergraduate degree in a maths-based subject, and teach the basics of markets, securities and investments, along with the mathematical and statistical methods that are used to track and model financial systems.

Gradwohl would advise anyone who develops an interest in finance as an undergraduate or earlier to enrol in a financial-engineering postgraduate programme rather than embarking on a PhD in physics or another ‘hard’ science. However, those who decide to enter the financial world having already earned a science PhD that gives them deep mathematical skills might well have a competitive edge in certain positions. Lubarsky notes that firms that concentrate on risk analysis and management usually prefer graduates of financial-engineering programmes, who learn relevant statistical and econometric techniques, whereas banks or companies looking



The Charging Bull statue on Wall Street in New York.

to devise complex derivatives are more likely to want people with science PhDs. Jobseekers with a science background should at the very least read several books about what quantitative analysis involves and learn the basics of investment before approaching potential employers. Those who are knowledgeable and persuasive in interviews don’t necessarily have to “tick all the boxes” in terms of formal qualifications, says Perraudin.

FINANCIAL INCENTIVE

For people entering the industry from academia, the salary of a quantitative analyst is a welcome change. A new hire with a PhD can expect to earn between US\$85,000 and \$110,000 a year, says Lubarsky, and could double that with bonuses; those with financial master’s degrees should expect to start at roughly \$65,000. But there are potential constraints. Locations are limited: the bulk of the work has traditionally been in New York and London, with smaller and less developed outposts in Tokyo, Hong Kong, Singapore and some other Asian cities. But that picture has begun to change as the job market diversifies, says Gradwohl. In the 1990s, quantitative analysts almost all worked on the ‘sell’ side of the financial world, analysing complex derivatives for companies that recommend and sell securities to investors. But today, many are on the more geographically widespread ‘buy’ side at hedge funds or asset-management firms, where they use their analytical skills to evaluate investments for companies that manage investor portfolios, says Gradwohl.

The work itself can be narrow and limiting at first: junior quantitative analysts spend their time using existing financial-analysis models for routine applications. “The standard model is not scientifically accurate, but it’s the model they use,” says an analyst who moved into the field from physics and is now working in Asia (his company required anonymity). He no longer has the freedom to choose how he spends his time on different projects, as he

had in academia. “You do what people want, you do it on time, and you do it right,” he says. Nevertheless, there are plenty of long-term prospects that offer some independence on the buy side or in management, for example, for those who are enthusiastic and adaptable: “If people like you, you can do whatever you want.”

Hodgdon spent only a few years in finance before deciding to move on. She got tired of working in New York and had no desire to embrace a lifestyle in which people spent their large salaries nearly as quickly as they earned them. “I don’t regret it, but I wouldn’t do it again,” she says. She moved away from the city and took a variety of other jobs — for which having Goldman Sachs on her résumé was a decided advantage — before setting up her own software company. Only those who genuinely like finance should pursue the profession, says Gradwohl: anyone who takes a finance job as a last resort or purely for the salary isn’t likely to thrive. Perraudin agrees that the working environment doesn’t suit everyone. “You’re surrounded by hard-charging, aggressive people,” he notes. “A large bunch of people get sick of it and disappear.”

But for those who take to it, finance can be exciting and interesting. When he first started out, Gradwohl worked on the buy side for several years, managing portfolios and monitoring investment performance; he received “immediate gratification” from seeing whether he was outperforming the competition. He has also come to appreciate the enormous and varied intellectual challenge of his work. Being in finance has broadened his horizons: Gradwohl’s work has touched on everything from game theory to sociology. “I couldn’t do it without physics, but physics wasn’t enough,” he says.

The economic upheavals that began in late 2008, driven by trading in complex derivatives chiefly based on housing markets, have changed the nature of finance, but have not diminished the need for quantitative analysis. Many of the exotic instruments most closely linked to the crisis — from ‘collateralized debt obligations’ to ‘credit-default swaps’ — have been swept away by the market and regulators, says Perraudin. But the urge to make money remains. Banks and hedge funds are continuing to explore new investments and securities. Extra regulation has, in fact, boosted the demand for risk analysis and management, he says.

Skilled quantitative analysts understand the risk of systemic failure and the limitations of quantitative models, suggesting that they will be in demand for the foreseeable future. “You can’t teach a smart finance person quantitative skills,” says Gradwohl, “but a smart scientist can learn finance.” ■

David Lindley is a freelance writer in Alexandria, Virginia.



“You can’t teach a finance person quantitative skills, but a scientist can learn finance.”

Ben-Ami Gradwohl

TURNING POINT

Andrew Dessler

Andrew Dessler, an atmospheric scientist at Texas A&M University in College Station, was named one of 21 Google Science Communication Fellows on 21 February.

What attracted you to climate research?

After my BA in physics at Rice University in Houston, I was sick of school. I worked for an investment bank on Wall Street for two years before realizing that I wasn't motivated by money — instead, I wanted to work on interesting problems. My father, an academic, suggested that the environment was going to be rife with difficulties in the future. Scientists had just discovered the ozone hole, so I worked on stratospheric ozone at the University of Maryland in College Park until the mid-1990s, when it was clear that the problem was essentially solved by the Montreal Protocol — the international agreement banning ozone-depleting chemicals.

What did you then focus your research on?

I looked for a big problem where I could make a contribution. I had worked on stratospheric water vapour previously, and moved down through the atmosphere to the troposphere to focus on the role of water vapour in climate. It took three to five years to make the transition to publishing entirely on climate research.

How and why did you enter the policy arena?

I heard that the White House Office of Science and Technology Policy (OSTP) was looking for someone. I had no idea what it would be like, but thought it would be a valuable experience — even if it slowed down my research a bit. I was there for the last year of the Clinton administration. Seeing how people at that level of government consume science made an impression on me. For better or worse, complex scientific results often get stripped of nuance by politicians. What's left — for example, long-term data trends — is what helps to shape policy.

How did you begin your efforts to communicate science to the public?

I helped teach a class at the University of Maryland on climate and policy after my stint at the OSTP. I noticed that the students perked up more when we discussed policy rather than black-body radiation. I co-wrote the book *The Science and Politics of Global Climate Change: A Guide to the Debate* with Edward Parson, now a professor at University of Michigan Law School in Ann Arbor. I also started a blog in 2006 for *Grist* magazine. At first, I was



enamoured with blogging, until I realized how repetitive it was to keep answering the same questions. I decided I wanted a more high-impact way to spend my time.

What did you think of 'Climategate' — the release of hundreds of leaked e-mails from the University of East Anglia in Britain?

It was a disaster. From a public debate standpoint, it was a huge setback. The University of East Anglia has been exonerated of any wrongdoing. But regardless, theirs is only a small part of the evidence supporting climate-change theory. Unfortunately, it is the kind of thing that is going to be with us for a long time.

Are you uneasy about stepping into the fray?

No. I'm convinced that the risks of climate change are severe. I view it as a moral responsibility to communicate those risks in the same way that if a train is heading for a small child, I've got to run and pull the child off the tracks. I'm not one to chain myself to a bulldozer, but I'm doing as much as I can with my personality and constitution.

Have you hopes for the Google Fellowship?

I realize I can't change the climate-change debate by myself. I need to find force multipliers — things that allow a small number of climate-change researchers to have an effect. For sceptics, spreading misinformation is a full-time job. But my full-time job is teaching and research, and I have to communicate in my spare time. There aren't many climate scientists in the world. We need ways to allow a few people to articulate what we know and the trade-offs that society faces. I'm hopeful that Google will find technological innovations that will help to spread the word. ■

INTERVIEW BY VIRGINIA GEWIN

ONLINE GAMES

Research role play

A free online game that simulates health-research projects in a virtual world aims to give players a taste of what it is like to be a scientist. In *Power of Research* (www.powerofresearch.eu), launched on 23 February by Biolution, a life-sciences consultancy in Vienna, players imitate the life of a researcher by choosing an institute and research topic, preparing for experiments, applying for funding, seeking staff, publishing results, attending conferences and collaborating with others. They need to be good at negotiation and time management. As players improve, they can win awards or become an institute leader. The game was developed with a €617,000 (US\$849,000) grant from the European Commission (EC), and is based on actual institutes and projects in the EC's research funding programme.

UNITED STATES

Rise in female scientists

The proportion of women occupying academic positions in biological and life sciences in the United States has risen slightly since 2001, finds a report by the US National Science Foundation (NSF). *Women, Minorities and Persons with Disabilities in Science and Engineering: 2011* says that women made up 31% of life scientists in 2001, and 36% in 2008. The growth matches the increasing number of women earning doctorates in the field, says Bobbie Mixon, an NSF spokesman. The report, released on 28 February, also found that by 2008, the overall number of life scientists working in industry and academia was 16% less than at its 23-year peak in 2006. Mixon links the decline to pharmaceutical and biotechnology layoffs.

UNITED KINGDOM

Opportunities in wheat

A £7-million (US\$11.4-million) research project into the genetic factors that affect wheat yields, plant size and resistance to drought and pests is hiring 17 researchers, including 8 postdocs, to work in molecular genetics, phenotyping and bioinformatics at centres around Britain. The project is the first of its kind in the country for more than 20 years, and is one of several worldwide. Researchers will build a database of genetic markers that can be used to create wheat varieties. The Biotechnology and Biological Sciences Research Council will fund three-year grants from April 2011, with the potential for extension.